

ISSN 2237-2083

REVISTA DE ESTUDOS DA LINGUAGEM

Faculdade de Letras da UFMG

JUL./SET. 2024

V. 32 – N. 3



LREVISTA DE ESTUDOS DA
LINGUAGEM

Universidade Federal de Minas Gerais

REITORA: Sandra Regina Goulart Almeida; VICE-REITOR: Alessandro Fernandes Moreira

Faculdade de Letras

DIRETORA: Sueli Maria Coelho; VICE-DIRETOR: Georg Otte

Editora-chefe

Janayna Maria da Rocha Carvalho (UFMG)

Editoras-associadas

Ana Regina Vaz Calindro (UFRJ)

Maria Mendes Cantoni (UFMG)

Conselho Editorial

Alejandra Vitale (UBA, Ciudad Autónoma de Buenos Aires, Argentina), Didier Demolin (Université de la Sorbonne Nouvelle Paris 3, Paris, França), Ieda Maria Alves (USP, São Paulo/SP, Brasil), Jairo Nunes (USP, São Paulo/SP, Brasil), Scott Schwenter (OSU, Columbus, Ohio, Estados Unidos), Shlomo Izre'el (TAU, Tel Aviv, Israel), Stefan Gries (UCSB, Santa Barbara/CA, Estados Unidos), Teresa Lino (NOVA, Lisboa, Portugal), Tjerk Hagemeijer (ULisboa, Lisboa, Portugal)

Editor de Arte

Emerson Eller

Projeto Gráfico

Stéphanie Paes

Secretaria

Lilian Souza dos Anjos, Ludmila Cunha

Revisão e normalização

Ana Regina Vaz Calindro (UFRJ), Janayna Maria da Rocha Carvalho (UFMG, Maria Mendes Cantoni (UFMG)

Diagramação

Kathleen Oliveira, Luísa Rocha Vasconcelos

REVISTA DE ESTUDOS DA LINGUAGEM





This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Os conceitos emitidos em artigos assinados são de responsabilidade exclusiva de seus autores.

Indexadores

Diadorim [Brazil]
DOAJ (Directory of Open Access Journals) [Sweden]
DRJI (Directory of Research Journals Indexing) [India]
EBSCO [USA]
EuroPub [England]
JournalSeek [USA]
Latindex [Mexico]
Linguistics & Language Behavior Abstracts [USA]
MIAR (Matriu d'Informació per a l'Anàlisi de Revistes) [Spain]
MLA Bibliography [USA]
OAJI (Open Academic Journals Index) [Russian Federation]
Portal CAPES [Brazil]
REDIB (Red Iberoamericana de Innovación y Conocimiento Científico) [Spain]
Sindex (Scientific Indexing Services) [USA]
Web of Science [USA]
WorldCat / OCLC (Online Computer Library Center) [USA]
ZDB (Elektronische Zeitschriftenbibliothek) [Germany]

REVISTA DE ESTUDOS DA LINGUAGEM, v. 1 - 1992 - Belo Horizonte, MG,
Faculdade de Letras da UFMG

Histórico:

1992 ano 1, n.1 (jul/dez)
1993 ano 2, n.2 (jan/jun)
1994 Publicação interrompida
1995 ano 4, n.3 (jan/jun); ano 4, n.3, v.2 (jul/dez)
1996 ano 5, n.4, v.1 (jan/jun); ano 5, n.4, v.2; ano 5, n. esp.
1997 ano 6, n.5, v.1 (jan/jun)

Nova Numeração:

1997 v.6, n.2 (jul/dez)
1998 v.7, n.1 (jan/jun)
1998 v.7, n.2 (jul/dez)

1. Linguagem - Periódicos I. Faculdade de Letras da UFMG, Ed.

CDD: 401.05

Faculdade de Letras da UFMG
Seção de Periódicos, sala 2017
Av. Pres. Antônio Carlos, 6627 – Pampulha
31270-901 – Belo Horizonte, MG – Brasil
Tel.: (31) 3409-6009
www.letras.ufmg.br/periodicos
periodicosfaleufmg@gmail.com

Sumário

- 677** O contexto de enunciados de popularização científica sobre o meio ambiente no Brasil
The Context of Scientific Popularization Utterances About the Environment in Brazil
Sheila Vieira de Camargo Grillo
- 700** Parâmetros vocais e prosódicos na fala de influenciadores digitais transgênero da mídia social brasileira
Vocal and Prosodic Parameters on Transgender Digital Influencers' Speech on Brazilian Social Media
Geovana Soncin; Eryne Alves Bafum; Gabriela Aparecida Rodrigues Gonçalves; Giovanna Caroline Borges; Karoline Araujo dos Santos; Eliana Maria Gradim Fabbron
- 738** O redobro do sujeito ele tem duas realizações prosódicas: o fraseamento prosódico das construções de redobro do sujeito na fala do Rio de Janeiro
Subject Doubling it has two Prosodic Realizations: Prosodic Phrasing of the Subject Doubling Constructions in the Speech of Rio de Janeiro
Eduardo Patrick Rezende dos Reis
- 762** Ferramentas analíticas para uma historiografia dos modelos sintáticos: rede taxonômica e glossário de metatermos da Grammatica da lingua Portuguesa (1540), de João de Barros
Analytical Tools for a Historiography of Syntactic Models: Taxonomic Network and Glossary of Metaterms from 'Grammatica da lingua Portuguesa' (1540) by João de Barros
Francisco Eduardo Vieira

- 804** Construção de corpus de 80 verbos de causalidade implícita no Português Brasileiro
Construction of a Corpus of 80 Implicit Causality Verbs in Brazilian Portuguese
Rute da Silva Barbalho; Renata Sabrinne Souza de Carvalho; Mahayana Cristina Godoy
- 824** Guiné Equatorial: política linguística, manutenção e obsolescência das línguas oficiais, étnicas e crioulas em um contexto ibero-africano
Equatorial Guinea: Language Policy, Maintenance and Obsolescence of Official, Ethnic, and Creole Languages in an Ibero-African Context
Gabriel Antunes de Araujo; Ana Lívia Agostinho
- 851** Metaphorical Phraseologies in a Learner Corpus: Investigating Translations from Brazilian Portuguese to English
Fraseologismos metafóricos em um corpus de aprendizes: investigando traduções do português brasileiro para o inglês
Jean Michel Pimentel Rocha
- 876** A Variação da Ordem Pronominal em Complexos Verbais na Variedade Moçambicana do Português
The Variation of Pronominal Order in Verb Complexes in the Mozambican Variety of Portuguese
Bento Orlando Mutoba; Norma Lucia Fernandes de Almeida
- 896** Demostración matemática: género discursivo y conexiones lógicas desde una mirada lingüística
Mathematical Proof: Discourse Genre and Logical Connexions from a Linguistic Perspective
Natalia Leiva Salum; Margarita Vidal Lizama

O contexto de enunciados de popularização científica sobre o meio ambiente no Brasil

The Context of Scientific Popularization Utterances About the Environment in Brazil

Sheila Vieira de Camargo Grillo
Universidade de São Paulo (USP) | São Paulo
SP | BR | CNPq
sheilagrillo@usp.br
<https://orcid.org/0000-0003-0480-2660>

Resumo: O objetivo é investigar como o contexto político e ideológico dos anos 2021-2022 no Brasil orientou a construção composicional, o conteúdo temático e o estilo de enunciados de popularização científica sobre o meio ambiente mediante as avaliações sociais e as ênfases valorativas de autores jornalistas e de discursos citados de cientistas. A atenção social recebida pelo tema do meio ambiente e a relevância do conceito de contexto para uma adequada abordagem do enunciado justificam a proposta. Para isso, foi reunido um *corpus* de enunciados da revista *Pesquisa FAPESP* publicados entre janeiro de 2021 e dezembro de 2022, totalizando 24 edições. Os procedimentos metodológicos de análise foram: caracterizar as esferas de produção, circulação e leitura dos enunciados; identificar os temas mais recorrentes; apontar como o contexto político e ideológico orientou a parte verbal dos enunciados do *corpus*; analisar o estilo em relação com a descrição das formas gramaticais. As teorias de M. Bakhtin, de P. Medvídev e de V. Volóchinov a respeito das relações entre as partes verbal e extraverbal do enunciado. As teorias de D. Mangueneau, F. Rastier e T. van Dijk evidenciaram a atualidade da discussão sobre o conceito de contexto. Síntese das descobertas: predomínio da modificação “modelo discurso direto preparado”; conexão entre a parte verbal e extra-verbal dos enunciados analisados por meio das avaliações sociais e das ênfases valorativas dos autores; ênfases valorativas de ameaça, risco aos biomas brasileiros e à saúde da população.

Palavras-chave: contexto; enunciado; popularização científica; environment.



Abstract: The objective is to investigate how the political and ideological context of the years 2021-2022 in Brazil guided the compositional construction, the thematic content and the style of scientific popularization utterances about the environment through social evaluations and the evaluative emphases of journalistic authors and from quoted speeches by scientists. The social attention given to the theme of the environment and the relevance of the concept of context for an adequate approach to the utterance justify the proposal. To this end, a *corpus* of utterances from the *Pesquisa FAPESP* magazine published between January 2021 and December 2022 was gathered, totaling 24 editions. The methodological analysis procedures were: characterize the spheres of production, circulation and reading of the statements; identify the most recurring themes; point out how the political and ideological context guided the verbal part of the statements in the *corpus*; analyze style in relation to the description of grammatical forms. The main theoretical basis is M. Bakhtin's, P. Medvedev's and V. Voloshinov's work regarding the relationships between the verbal and extraverbal parts of the utterance. The theories of D. Maingueneau, F. Rastier and T. van Dijk highlighted the relevance of the discussion on the concept of context. Main findings: predominance of the "prepared direct speech model" modification; connection between the verbal and extra-verbal parts of the utterances analyzed through social evaluations and the authors' evaluative emphases; evaluative emphases of threat and risk to Brazilian biomes and the health of the population.

Keywords: context; utterance; scientific popularization; environment.

1 Introdução

À luz de teorias contemporâneas sobre o contexto (Maingueneau, 1993; Rastier, 1998; Van Dijk, 2012[2007]) e sobretudo dos conceitos de contexto presentes nos trabalhos de Mikhail Bakhtin, Pável Medviédev e Valentín Volóchinov dos anos 1920, o objetivo deste artigo é: *investigar como o contexto político e ideológico dos anos 2021-2022 no Brasil orienta a construção composicional, o conteúdo temático e o estilo de enunciados de popularização científica sobre o meio ambiente mediante as avaliações sociais e as ênfases valorativas de autores jornalistas e de discursos citados de cientistas.*

Em outros termos, trata-se de mostrar como o contexto exterior do enunciado se torna interior a ele ao direcionar a seleção e a organização de seus elementos composticionais, temáticos e estilísticos. A divulgação ou popularização científica é aqui concebida como um tipo de relação dialógica – no sentido bakhtiniano (Bakhtin, 2022[1929]) de relações semânticas – entre enunciados da esfera científica e de outras esferas da atividade humana (Grillo, 2013).

O *corpus* da pesquisa são reportagens sobre o meio ambiente publicadas pela “Fundação de Amparo à Pesquisa do Estado de São Paulo” (FAPESP) nas edições mensais da revista de divulgação ou popularização científica *Pesquisa FAPESP*, entre janeiro de 2021 e dezembro de 2022, totalizando 24 números. Esse período foi escolhido em razão de a proximidade das eleições presidenciais ter provocado uma intensificação da polarização ideológica e política no Brasil com impacto nos enunciados sobre o meio ambiente. A escolha da revista *Pesquisa FAPESP* se justifica por ser uma publicação da comunidade científica brasileira do Estado de São Paulo, cujo governo à época fez oposição à política do governo federal. O meio ambiente foi selecionado por receber uma atenção especial da sociedade brasileira e internacional contemporânea, tornando-se objeto de disputa na esfera política brasileira e tema privilegiado para investigar as relações entre as partes verbal e extraverbal do enunciado.

2 A atualidade e a importância epistemológica do contexto para as teorias do discurso¹

Como ponto de partida, três teorizações do contexto são desenvolvidas por três eminentes linguistas do texto e do discurso: a análise do discurso de Dominique Maingueneau, a semântica interpretativa de François Rastier e a análise do discurso crítica (ADC) de Teun van Dijk. A escolha das teorias desses autores se justifica pelo fato de: eles terem formulado reflexões explícitas e amplas da noção de contexto; o contexto ser um elemento constituinte de suas abordagens do discurso (Maingueneau), do texto (Rastier) e do discurso/texto (Van Dijk); as teorizações terem sido elaboradas no campo da linguística ou dos estudos da linguagem, área de nossa atuação e interesse. A exposição de trabalhos de linguistas contemporâneos objetiva revelar a atualidade e a relevância da noção de contexto no horizonte dos estudos da linguagem, bem como permite melhor avaliar, pela comparação, as contribuições e as especificidades do método sociológico desenvolvido por Mikhail Bakhtin, Pável Medvídev e Valentín Volóchinov nos anos 1920, nossa principal base teórica.

Começamos pelo linguista e analista do discurso francês Dominique Maingueneau, que, no início dos anos 1990, aponta uma nova concepção do fato literário como “um ato de comunicação no qual o dito e o dizer, o texto e seu contexto são indissociáveis”² (1993, p. VI)³. Maingueneau se propõe a tratar do “entorno imediato do texto”⁴ (1993, p. 23) formado por quatro elementos:

¹ Parte da exposição teórica aqui apresentada foi objeto de um artigo já publicado (Grillo, 2023). Aqui utilizamos uma fração do que foi lá elaborado para demonstrar sua pertinência à análise dos enunciados do *corpus*

² un acte de communication dans lequel le dit et le dire, le texte et son contexte sont indissociables.

³ Embora *Le contexte de oeuvre littéraire. Énonciation, écrivain, société* (1993) tenha saído há 30 anos, esse texto foi escolhido pela fato de o contexto ser seu objeto central de reflexão.

⁴ abords immédiats du texte.

- ◆ Os modos de posicionamento do autor no campo literário, do qual Maingueneau destaca os gêneros do discurso, concebidos como um tipo de atividade social exercida em determinadas circunstâncias por protagonistas autorizados, e que, a depender do gênero, se inscrevem em percursos históricos variados dentro do campo;
- ◆ O suporte material do enunciado (a voz, a escrita, o impresso, a língua) compreendido por Maingueneau, em diálogo R. Debray (1991), como modos de transmissão e redes de comunicação constitutivos do sentido de uma obra ou enunciado⁵;
- ◆ A situação de enunciação pressuposta (condições de enunciação ligadas a cada gênero do discurso e, portanto, socialmente reconhecidas) e validada pelo enunciado por meio de sua cenografia (conjunto interconectado de suas referências subjetivas – enunciador e coenunciador –, temporal e espacial construídas pelo enunciado e inscritas nele). A cenografia está ligada ainda a maneiras de dizer que exprimem uma personalidade e uma corporalidade, ou seja, o ethos do enunciador na relação com o co-enunciador;
- ◆ Por fim, o último elemento é a enunciação no enunciado, ou seja, o quadro pragmático ou ato de enunciação ou ainda o ato de comunicação na relação com o “dito” ou mundo representado. Em outros termos, o ato de enunciação pelo qual o mundo é visto em um enunciado.

A obra de Bakhtin é referência constante de *Le contexte de oeuvre littéraire. Énonciation, écrivain, société* (1993), em passagens como: “conforme observou Bakhtin, o estudo do discurso se faz na fronteira entre disciplinas tradicionais”⁶ (p.18), “a importância da reflexão sobre a intertextualidade, muito viva desde o fim dos anos 70 e que faz eco ao dialogismo de Bakhtin”⁷ (p. 21); “a contracultura carnavalesca, que, por meio do escárnio, visava subverter a cultura oficial”⁸ (p. 36); “como afirma com bastante razão Bakhtin, a comunicação verbal supõe a existência de gêneros do discurso”⁹ (p. 65).

Em texto mais recente, D. Maingueneau em parceria com P. Charaudeau (2002) distinguem “situação de comunicação” e “situação de enunciação” do seguinte modo: a “situação de comunicação “seria o contexto efetivo de um discurso”¹⁰ (p. 229), no qual a enunciação se aproxima do ato de linguagem, e a situação de enunciação “seria um sistema de coordena-

⁵ Duas considerações sobre o suporte merecem nota: primeiramente, o modo de transmissão impresso ao separar o texto da oralidade, provocou o desaparecimento da “voz” do autor e possibilitou o surgimento de teorias literárias como o estruturalismo, que dissociou o texto do contexto e que seria inimaginável em sociedades com predomínio da literatura oral; e, em segundo lugar, o fato de o escritor literário ser confrontado a uma interação de línguas e usos, que Maingueneau chama de “interlíngua”, fazendo menção aos conceitos bakhtinianos de “heteroglossia”, “dialogismo” e “plurilinguismo”, este último concebido em sua dimensão externa (relação entre línguas) e interna (a diversidade dentro de uma mesma língua).

⁶ comme le notait déjà Bakhtine l'étude du discours se joue 'aux frontières' de disciplines traditionnelles.

⁷ la réflexion sur l'intertextualité, très vivante depuis la fin des années 70 et qui fait écho au 'dialogisme' de Bakhtine.

⁸ la contre-culture ‘carnavalesque’ qui par la dérision visait à subvertir la culture officielle

⁹ Comme l'affirme très justement M. Bakhtine, la communication verbale suppose l'existence de gêneros de discours.

¹⁰ serait le contexte effectif d'un discours

das abstratas associadas a toda produção verbal”¹¹ (p. 229). A enunciação é considerada um componente autônomo da teoria da linguagem (Greimès, Courtès, 2008[1993]) formado por estruturas de mediação que asseguram a colocação em discurso-enunciado do sistema social que é a língua. Charaudeau e Maingueneau (2002) enfatizam que não se trata da distinção entre o geral e o singular, pois existem invariantes tanto na “situação de comunicação” quanto na “situação de enunciação”. Conforme mostraremos na próxima seção, a “situação de comunicação” é mais próxima do conceito de contexto desenvolvido pelo Círculo de Bakhtin na segunda metade dos anos 1920.

O segundo teórico do contexto é François Rastier¹² (2022[1998]), para quem a teorização sobre o contexto evidencia aspectos de uma mudança epistemológica nas ciências da linguagem, com seu emprego crescente sobretudo na semântica e na pragmática. O contexto permite ainda opor duas grandes problemáticas ou tradições de reflexão sobre a linguagem: a tradição lógico-gramatical e a tradição retórico-hermenêutica (Rastier, 2022[1998]). Esta concebe o texto como um sistema pensado segundo a oposição ocorrência-fonte e retomada, e não tipo-ocorrência. Essa problemática se ocupa do contínuo, segundo o princípio hermenêutico de que o global condiciona e determina o local¹³ (uma passagem ou uma unidade discreta). Nessa orientação, recontextualizar a linguagem impõe a restauração da intersubjetividade – distinguindo a situação de interpretação e a situação de enunciação, cada uma com seu universo de referências, de assunções e de responsabilidades – e a história – presente sob a forma da intertextualidade contemporânea e passada, ou seja, o contexto não é só o aqui e agora, mas ultrapassa a situação. O conceito de gênero permite conectar o contexto linguístico e a situação, pois ele é um princípio organizador do texto e um modo semiótico da prática em curso. Além disso, todo texto é interpretado no interior de um *corpus* formado, em primeiro lugar, pelos textos que pertencem a um mesmo gênero. Para Rastier (2022[1998]), o texto é a unidade linguística fundamental, cuja análise orienta o acesso às unidades de nível inferior e cuja unidade superior é o *corpus*.

Passamos ao conhecido livro *Discurso e Contexto. Uma abordagem sociocognitiva* (2012[2007]) do linguista Teun Van Dijk. Nele, o autor faz uma extensa revisão bibliográfica da noção de contexto em ciências humanas e, assim como Rastier, Van Dijk também vê a reflexão sobre o contexto provocar uma mudança epistemológica nos estudos da língua e do discurso: “o contextualismo implica que os fenômenos precisam ser estudados em relação a uma situação ou entorno” (Van Dijk, 2012[2007], p. 28), o que contrasta com teorias “descontextualizadas, abstratas, estruturalistas, formalistas, autônomas” (Van Dijk, 2012[2007], p. 28) sobre os fenômenos da linguagem. Van Dijk considera que, depois da Segunda Guerra Mundial, teorias formalistas, estruturalistas ou autônomas dominaram a esfera científica da linguística e das ciências humanas. Sempre segundo Van Dijk, foi só nos anos 1970 e 1980, em disciplinas como a Etnografia da Fala, a Pragmática, a Sociolinguística e a Análise do Discurso

¹¹ serait un système de coordonnées abstraites associées à toute production verbale

¹² Rastier considera que Bakhtin retomou, em seu dialogismo, os primeiros românticos (sobretudo Schleiermacher) e se situa, apesar de sua “coloração” marxista, na tradição do idealismo subjetivo.

¹³ Considerado o pai da hermenêutica moderna, Schleiermacher, aliando o princípio da composicionalidade ao da globalidade, postula: “não somente a compreensão do todo é condicionada pela do detalhe, mas inversamente a compreensão do detalhe é determinada pela compreensão do todo” [non seulement la compréhension du tout est conditionnée par celle du détail, mais encore inversement la compréhension du détail est déterminée par la compréhension du tout] (1987 [1809-1810], p. 77)

Crítica, que se desenvolveu uma “abordagem integrada do uso linguístico e dos eventos comunicativos voltada para o contexto e nele inserida” (Van Dijk, 2012[2007], p. 300-301).

Como a própria designação evidencia, Van Dijk assenta sua abordagem sociocognitiva, por um lado, em uma base cognitivista (representações mentais¹⁴ subjetivas), e, por outro, em uma base social (conhecimentos, atitudes, ideologias, gramática, regras, normas e valores socialmente compartilhados por comunidades discursivas) para propor que a produção e a interpretação dos discursos ocorre sob o controle de modelos mentais de contextos, ou seja, “não um tipo de situação social objetiva, e sim construtos dos participantes, subjetivos embora socialmente fundamentados, a respeito das propriedades que para eles são relevantes em tal situação, isto é, modelos mentais” (Van Dijk, 2012[2007], p. 87)¹⁵. Van Dijk enfatiza serem esses modelos mentais flexíveis e negociarem os falantes constantemente sua interpretação dos aspectos relevantes da situação comunicativa. A relação entre os modelos mentais dinâmicos de contexto e as estruturas de discurso é de *controle* das possíveis variações da língua, do texto e do conhecimento.

3 O contexto no método sociológico desenvolvido por Bakhtin, Medviédev e Volóchinov

Diante da grande variedade e complexidade do conceito de contexto, constatada a partir de uma pesquisa no conjunto da obra de Mikhail Bakhtin e de Valentín Volóchinov, bem como no livro *O método formal nos estudos literários. Introdução crítica a uma poética sociológica* (1928) de Pável Medviédev, abordaremos, de modo resumido, apenas as obras relativas ao método sociológico. Esse método foi desenvolvido durante o período de inserção no ou proximidade desses autores ao ILIAZV (Instituto da História Comparada das Literaturas e das Línguas do Ocidente e do Oriente) em Leningrado, de 1925 a 1930 (Grillo; Américo, 2019), por considerar que as formulações desse período são mais produtivas à análise do *corpus* desta pesquisa.

Nossa investigação bibliográfica seguirá a cronologia de publicação dos textos e incidirá sobre aqueles artigos de Valentin Volóchinov, presentes na coletânea *A palavra na vida e a palavra na poesia: ensaios, artigos, resenhas e poemas* (2019), com as formulações mais elaboradas sobre a noção de contexto, bem como sobre as obras de maior fôlego publicadas pelos três pesquisadores, a saber: *O método formal nos estudos literários. Introdução crítica a uma poética sociológica* (2012[1928]) de Pável Medviédev, *Marxismo e filosofia da linguagem. Problemas fundamentais do método sociológico na ciência da linguagem* (2021[1929]) de Valentin Volóchinov e *Problemas da obra de Dostoiévski* (2022[1929]) de Mikhail Bakhtin.

No artigo *A palavra na vida e a palavra na poesia. Para uma poética sociológica* (2019[1926]), o objetivo geral é mostrar que o método sociológico é capaz de superar a ruptura entre forma e conteúdo, teoria e história, poética teórica (que trata de forma, tema e estilo poéticos) e poética sociológica (que aborda a influência do meio social extra-artístico sobre as formas

¹⁴ O conceito de “representação” nomeia o modo original da presença de um objeto externo no espírito humano, ou seja, um objeto externo físico é “presente de novo” (re-presentado), mas sobre uma outra forma (Meunier, 2002).

¹⁵ No âmbito de uma semiótica da interpretação agregadora das ciências da cultura, Bouquet (2002) fala em **cognição situada**, ou seja, uma cognição situada em um quadro cultural.

literárias). A arte literária é um dos campos da criação ideológica (ao lado do jurídico, do cognitivo etc.): “*uma forma específica de comunicação social, realizada e fixada no material da forma artística*” (Volóchinov, 2019[1926], p. 116). Dessa perspectiva, Volóchinov propõe o objetivo de “compreender a forma do enunciado poético enquanto a forma dessa comunicação estética específica, realizada no material da palavra” (Volóchinov, 2019[1926], p. 117).

A partir desse propósito, Valentín Volóchinov assevera que é no processo da comunicação social que a palavra no enunciado artístico ganha vida, sendo o contexto extraverbal um elemento integrante tanto do conceito de enunciado literário quanto do *não literário por meio do percebido, compreendido e avaliado pelo falante, pelo personagem/acontecimento e pelos ouvintes*. Nesses enunciados são as avaliações sociais (posições ideológicas constitutivas do horizonte ideológico amplo de uma época e de uma sociedade) que conectam a parte verbal com o contexto do enunciado. A posição ativa dos falantes sob a influência dessas avaliações sociais elabora ênfases valorativas que inscrevem o contexto extraverbal na parte verbal do enunciado.

Em *O método formal nos estudos literários. Introdução crítica a uma poética sociológica* (2012[1928]), Pável Medviédev formula sua poética sociológica em diálogo especialmente com os primeiros escritos dos formalistas russos, produzidos entre 1914 e 1919. Logo de início, o autor explicita que a questão norteadora do livro é superar a ruptura entre o estudo de um fenômeno ou produto ideológico (uma obra de arte, um trabalho científico, uma cerimônia religiosa compostos por seus materiais, formas e propósitos) e as especificidades dos campos da criação ideológica.

A resposta a essa questão é orientada pela interpretação dialética¹⁶ de que os fatores sociais externos influenciam a natureza interna da literatura (seu enredo, estilo, composição etc.), tornando-se internos. Assim como em V. Volóchinov, a parte material (verbal, sonora, plástica etc.) e extraverbal ou contexto se unem na composição do enunciado mediados pela avaliação social. Para P. Medviédev, o contexto sociohistórico do enunciado ou de qualquer fenômeno ideológico é formado pela situação comunicativa imediata, pelo campo ou esfera ideológico/a específico/a do enunciado (literário, artístico, científico etc.) e pelo horizonte ideológico geral de uma época e de uma sociedade (conjunto dos objetos-signos – obras de arte, símbolos religiosos, afirmações científicas etc. – que constituirão a consciência social de uma coletividade). A relação entre o enunciado e o contexto é de reflexo e de refração, ou seja, o enunciado não é nem totalmente autônomo nem submisso por completo ao contexto. Essa compreensão orienta a metodologia de isolamento do objeto de pesquisa do método sociológico: os fenômenos ideológicos (obras de arte, trabalhos científicos etc.) são sócio-históricos, ou seja, não podem prescindir do contexto.

Em 1929, Mikhail Bakhtin publica a primeira versão de seu conhecido livro sobre Fiódor Dostoiévski, *Problemas da obra de Dostoiévski* (2022[1929]). Em seu prefácio, o autor se propõe, do ponto de vista metodológico, a superar a separação entre uma análise ideológica e uma análise formal a fim de encontrar “a ideologia que determinou sua forma artística” (Bakhtin, 2022[1929], p. 52) e, para isso, explicita sua convicção de que:

¹⁶ Spinelli (2008) argumenta que o crítico literário e ensaísta Antônio Cândido propõe uma interpretação dialética, ou seja, o elemento externo se torna interno à obra, do romance “Senhora” (1874), de José de Alencar. Este assimila a dinâmica da exploração econômica à temática romântica, em uma narrativa em que um casamento mediado por manobras e concessões expressa a “mineralização da personalidade, tocada pela desumanização capitalista” (Spinelli, 2008, p. 46).

toda criação literária é interna e imanemente sociológica. Nela cruzam-se forças sociais vivas, cada elemento de sua forma está permeado por avaliações sociais vivas. É por isso que uma análise puramente formal deve tomar cada elemento da estrutura artística como um ponto de refração das forças sociais vivas (...) (Bakhtin, 2022[1929], p. 52)

Há, aqui, uma grande coerência entre os textos de Bakhtin, Medvídev e Volóchinov analisados até o momento: a forma artística (objeto de estudo do autor, mas poderia ser a forma científica, a jornalística, a religiosa etc.) não pode ser compreendida sem as forças sociais nela refratadas, ambas mediadas pelas avaliações sociais. Na análise do romance polifônico, Bakhtin sustenta que as teses ideológicas extraídas do contexto social, histórico e ideológico de Dostoiévski são refratadas no conteúdo do romance mediadas pela “arquitetônica artística das obras de Dostoiévski” (Bakhtin, 2022[1929], p. 60); esta também chamada de forma, visão ou vontade artística, ou ainda da “ideologia constituinte da forma” (Bakhtin, 2022[1929], p. 133).

No decorrer do livro, o contexto extraverbal é *composto por vários aspectos da obra literária: a orientação da palavra autoral para seu objeto* (Bakhtin, 2022[1929], p.174); as “intenções do autor” que podem ou não ser expressas diretamente na narração do romance (Bakhtin, 2022[1929], p.166); as “avaliações sociais segmentadas” refletidas e refratadas no estilo do autor (Bakhtin, 2022[1929], p.168); a antecipação das objeções, avaliações e pontos de vista do interlocutor na palavra literária (Bakhtin, 2022[1929], p.176); a orientação autoral à palavra alheia, seja na mesma direção semântica, seja em direções semânticas distintas como na paródia (nos limites das intenções do autor) e na polêmica (influenciando a palavra do autor de fora). Ao defender que a orientação à ou relação com a palavra alheia é um dos principais traços do romance de Dostoiévski, Mikhail Bakhtin afirma a natureza social interna da linguagem enquanto “meio de comunicação social em eterno movimento” (Bakhtin, 2022[1929], p.184), pois a “vida da palavra está na passagem de uma boca a outra boca, de um contexto a outro contexto, de uma coletividade social a outra, de uma geração a outra geração” (Bakhtin, 2022[1929], p.183-184, grifo nosso).

Em *Marxismo e filosofia da linguagem. Problemas fundamentais do método sociológico na ciência da linguagem* (2021[1929]), Valentin Volóchinov objetiva (mencionando o interesse dos formalistas pela palavra) investigar o lugar da linguagem no pensamento marxista e, para isso, propõe as seguintes questões centrais: a realidade concreta dos fenômenos linguísticos, “o papel produtivo e a natureza social do enunciado” (Volóchinov, 2021[1929], p. 87) e o problema do enunciado alheio.

Nessa obra de V. Volóchinov, o contexto pode ser formado pelo entorno extraverbal – a situação comunicativa imediata (locutores e suas posições sociais, tempo, espaço etc.), o contexto sócio-ideológico amplo (esferas ideológicas, a mesma coletividade linguística, uma sociedade organizada de modo específico etc.), a orientação para o interlocutor, as avaliações sociais – e pelo entorno verbal – os enunciados-fonte e os enunciados-resposta de um enunciado, o discurso interior do falante, o discurso verbal (autoral ou do personagem) assimilador e transmissor do discurso alheio. Ao responder à pergunta sobre a natureza concreta da linguagem, Volóchinov defende que o contexto extraverbal é parte integrante do acontecimento social da interação discursiva, esta tomada como a essência da língua/linguagem, e elemento determinante do tema ou sentidos singulares dos enunciados em constante tensão com as significações mais estáveis da língua.

Por fim, abordaremos o contexto, de modo conjunto, nos três artigos que formam o projeto de popularização científica *A estilística do discurso literário* (1930): I. O que é a linguagem/língua?; II. A construção do enunciado; III A palavra e sua função social.

Em razão do propósito popularizador desses três artigos, Valentin Volóchinov didatiza, exemplifica, explicita, revisita os aspectos do contexto já trabalhados em *Marxismo e filosofia da linguagem. Problemas fundamentais do método sociológico na ciência da linguagem*. Quatro aspectos sobre o contexto nesses artigos merecerem destaque: primeiramente, a afirmação do contexto como “parte extraverbal do enunciado”, ou seja, não é um acréscimo, mas parte integrante; em segundo lugar, o enunciado contém uma parte verbal e outra não verbal, esta formada por “gestos”, “expressão do rosto”, “pose do corpo” (Volóchinov, 2019[1930]b, p. 281) e pelas “condições técnicas exteriores” (no caso de uma obra literária, a redação, a tipografia, o mercado editorial etc.) (Volóchinov, 2019[1930]a, p. 263-4)¹⁷; em terceiro lugar, o papel da língua na criação da cultura ou sistemas ideológicos constituídos, pois a língua – como forma materializada da comunicação social – exerce uma influência inversa sobre essas condições ao marcar uma “ruptura com o mundo da natureza” e ao criar o mundo da cultura ou dos sistemas ideológicos (ciência, arte, moral, direito etc.) (Volóchinov, 2019[1930]a, p. 248); por fim, a influência do contexto imediato e amplo na definição do tema, da entonação, da escolha das palavras e da disposição delas no todo do enunciado.

A partir desse breve percurso bibliográfico, destacamos os seguintes aspectos do contexto que nortearão a análise do nosso *corpus* de pesquisa:

- 1 Os enunciados, obras, discursos ou textos são constituídos por uma parte verbal e outra extraverbal (esferas ideológicas ou da atividade humana, horizonte ideológico de uma época ou sociedade, coenunciadores, tempo, espaço etc.), ou seja, o contexto é parte integrante e inalienável do enunciado;
- 2 O enunciado é formado ainda por aspectos verbais (conteúdo temático, estilo e construção composicional¹⁸) e não verbais (gestos, fotos, tipografia, condições técnicas exteriores etc.), os quais sofrem coerções e são influenciados pela parte extraverbal do enunciado (situação imediata, contexto sócio-histórico-ideológico, horizonte ideológico);
- 3 A parte extraverbal ou contexto do enunciado “é potencialmente inacabável” (Bakhtin, 2017[1970–71], p. 44), sendo constituída, entre outros, pela a situação comunicativa imediata, pelas esferas da atividade humana em que o enunciado é produzido, recebido e circula (científica, política, jornalística etc.), pelo horizonte ideológico geral de uma época e de uma sociedade (conjunto dos objetos-signos – obras de arte, símbolos religiosos, afirmações científicas etc. – que constituirão a consciência social de uma coletividade);
- 4 O contexto de um enunciado é também a cadeia de outros enunciados com os quais se relaciona, isto é, seu contexto verbal. Este é formado, por um lado, pelos enunciados

¹⁷ Vemos aqui que Volóchinov abordou em termos próximos o que mais tarde foi chamado por Maingueneau de suporte material do enunciado.

¹⁸ Consideramos que esses três conceitos presentes no texto “Os gêneros do discurso” de Mikhail Bakhtin (2016[1953-54]) já são suficientemente conhecidos da comunidade científica contemporânea, não necessitando, por isso, de maiores esclarecimentos.

sobre o mesmo tema ou do mesmo gênero; e, por outro, pelos discursos citados, relatados ou reportados. Em relação ao primeiro aspecto, Maingueneau (1993) destaca que, ao se inscrevem em determinado gênero, um discurso remete a percursos históricos variados dentro de um campo. Já Rastier (2022[1998]) concebe a relação entre textos segundo a oposição ocorrência-fonte e ocorrência-retomada em que a *língua* é um sistema em constante mudança devido a usos e dinâmicas históricos.

- 5 As avaliações sociais (posições ideológicas constitutivas do horizonte ideológico amplo de uma época e de uma sociedade) conectam o contexto do enunciado com a parte verbal. Já a posição ativa dos falantes sob a influência dessas avaliações sociais elabora ênfases valorativas que inscrevem o contexto extraverbal na parte verbal do enunciado.

4 Esferas da divulgação científica de Pesquisa FAPESP (2021-2022): orientações metodológicas

Entre janeiro de 2021 e dezembro de 2022, foram publicadas 24 edições da revista *Pesquisa FAPESP* (número 299 ao 322). A FAPESP tem como uma das atividades afins – segundo “Carta da Editora” (número 315, maio 2022) por ocasião da comemoração dos 60 anos da Fundação de Amparo à Pesquisa do Estado de São Paulo – a “difusão de conhecimento para a sociedade”. Com esse propósito, a revista *Pesquisa FAPESP* foi criada nos anos 1990 com o objetivo de “ampliar o acesso aos resultados de pesquisas científicas no Brasil, de tratar de processos, instituições e pessoas envolvidos nesse fazer científico” e de “viabilizar uma publicação jornalística que zela pela precisão sem perder de vista a perspectiva dos leitores” (grifos nossos). Com essas referências e dados históricos, realizamos o primeiro passo metodológico da pesquisa, ao inserir as 24 edições percorridas em suas esferas de produção, circulação e leitura. A revista *Pesquisa FAPESP* é o resultado, sobretudo, do diálogo entre as esferas científica, jornalística e a difusa ideologia do cotidiano (Volóchinov, 2021[1929]) da sociedade brasileira.

A esfera ou campo jornalística/o se faz presente na figura dos autores jornalistas (diretora de redação, editor-chefe, editores, repórteres, redatores etc.) e nos gêneros estruturadores dos números: fotolab, nota/notícia, carta dos leitores, carta da editora (editorial), reportagem, dossiê de capa, reportagem de capa, entrevista, obituário, perfil, entrevista. Esferas/campos e gêneros conectam: os enunciados a percursos históricos variados dentro do campo, segundo Maingueneau (1993); os textos à situação e aos textos pertencentes a um mesmo gênero, nos termos de François Rastier (1998 e 2022); os enunciados ao contexto sócio-histórico (situação comunicativa imediata, campo ideológico específico do enunciado e horizonte ideológico geral de uma época e de uma sociedade), na perspectiva de Pável Medvídev; os enunciados aos seus contextos verbais (os enunciados-fonte e enunciados-resposta do enunciado, o discurso autoral assimilador e transmissor do discurso alheio), segundo Mikhail Bakhtin e Valentín Volóchinov.

Ainda no âmbito desse primeiro passo metodológico realizou-se a leitura das 24 edições do *corpus* em busca de enunciados sobre o tema do meio-ambiente ou ecologia, os quais estavam presentes em todos os 24 números. Nas edições de *Pesquisa FAPESP* dos anos de 2021

e 2022, o aquecimento da Terra em razão da emissão de gases de efeito estufa (GEE) é um dos temas mais recorrentes nos diferentes gêneros, sendo a agropecuária brasileira uma emissora importante desses gases.

O segundo passo metodológico foi, à luz da questão central de pesquisa, mostrar a relação entre o enunciado e o contexto extraverbal, ou seja, apontar como o contexto político e ideológico dos anos 2021-2022 mediado pelas avaliações sociais e pelas ênfases valorativas dos autores jornalistas e dos discursos citados dos cientistas orientaram a construção composicional, o conteúdo temático e o estilo dos enunciados do *corpus*. Para isso, foram selecionadas duas reportagens do n. 320 de Outubro de 2022, em que o contexto político das eleições (1º. Turno em 02/10/2022 e 2º. Turno em 30/10/2022) para presidente se revelou determinante na seleção dos temas, dos discursos citados, do estilo, das imagens e da construção composicional. A análise desses dois conjuntos foi acompanhada do cotejo com fragmentos de enunciados de outras edições do *corpus*, a fim de mostrar a recorrência do conteúdo temático abordado.

Por fim, as **análises do estilo** – este concebido como a seleção dos recursos gramaticais, sintáticos e lexicais (Bakhtin, 2016[1953-54]) – se orientou pelas formulações tanto de Bakhtin (2013) quanto de Volóchinov (2021[1929], 2019[1930]), segundo os quais as formas gramaticais não podem ser estudadas sem a consideração do seu significado estilístico. Ambos os autores partem, sobretudo, de descrições de fenômenos sintáticos (período composto por subordinação sem conjunção ou orações subordinadas não-conjuntivas, formas de transmissão do discurso alheio, seleção e organização de palavras na frase) para explorar seus sentidos estilísticos. Foi esse procedimento metodológico que orientou a análise de aspectos estilísticos dos enunciados do *corpus* e motivou o uso da gramática de Neves (2000), obra reconhecida pela qualidade de sua abordagem científica da língua.

5 A natureza sob ameaça

A primeira reportagem¹⁹ selecionada tem o título “A aldeia dos primatas” e o título-auxiliar “Mais de 70% das espécies de macacos podem ser encontradas em terras indígenas” (Pesquisa FAPESP, outubro de 2022, ano 23, n. 320, p. 56-57). Destacamos dois aspectos nos quais esse enunciado reflete e refrata o contexto extraverbal: o **peritexto**²⁰ – títulos e fotografias – e as **ênfases valorativas** dos autores condicionadas pelas **avaliações sociais** que orientaram tanto o contexto autoral quanto o discurso alheio introduzido e assimilado no corpo enunciado.

¹⁹ Gênero da esfera jornalística que se caracteriza pela ampliação do tema ou acontecimento abordado mediante investigação e exposição das fontes (Grillo, 2004).

²⁰ Segundo Genette (1987), o peritexto é uma categoria espacial (“catégorie spatiale”, p. 10) composta pelos elementos que circundam de modo imediato o texto: título, prefácio, capa etc.

Imagen 1 – A aldeia dos primatas

ECOLOGIA

A ALDEIA DOS PRIMATAS

Mais de 70% das espécies de macacos podem ser encontradas em terras indígenas

Eduardo Geraque

A destruição das matas e das populações tradicionais coloca os povos indígenas em risco a sobreveleira das espécies de macacos do planeta. Segundo uma matéria publicada em agosto na revista *Science Advances*, 30% das áreas de distribuição de primatas não humanos existentes de terras indígenas e 77% de suas 521 espécies podem ser encontradas nesses territórios. Cerca de metade das espécies que vivem em terras indígenas do Neotrópico (América Central e do Sul), pouco mais de 100 espécies, são consideradas ameaçadas de extinção. Entre as que podem ser encontradas nos territórios dos povos tradicionais, a proporção cai para 55%. "Quanto maior for a área de distribuição protegida, menor é o risco de as espécies de primatas em declínio ficarem ameaçadas de extinção ou terem populações em declínio", afirma

encontram mais protegidas nos locais em que os povos tradicionais conseguiram manter sua cultura de forma mais autônoma em relação à sociedade atual. "Há uma sobreproteção, pelo lado da parte, da conservação das espécies com os territórios indígenas", diz o ecólogo Ricardo Dobrovolski, da Universidade Federal da Bahia (UFBA), um dos autores da matéria. "As espécies que vivem nesses focos de resistência têm uma chance menor de estar sob risco de extinção e formam algo que podemos chamar de quilombos da biodiversidade."

Mais de 90% das espécies cujo habitat é de terras indígenas são consideradas ameaçadas de extinção. Entre as que podem ser encontradas nos territórios dos povos tradicionais, a proporção cai para 55%. "Quanto maior for a área de distribuição protegida, menor é o risco de as espécies de primatas em declínio ficarem ameaçadas de extinção ou terem populações em declínio", afirma

Alejandro Estrada, do Instituto de Biología da Universidade Nacional Autônoma do México, primeiro autor do trabalho. Nos trópicos americanos e na África continental, 44% das espécies de primatas vivem em terras indígenas de declínio em comparação com o arredio. Seu estado de conservação é classificado como vulnerável, em perigo ou criticamente em perigo. A situação é muito mais grave na região indonésia e em Madagascar, onde 90% das espécies, eventualmente, estão ameaçadas de extinção. "O estado de conservação desse grupo animal difere significativamente entre as regiões do mundo", comenta o primatólogo Alan Gardner, da Universidade de Illinois, nos Estados Unidos, que também participou do trabalho.

A forma como uma determinada paisagem rica em biodiversidade é protegida pode ser relevante, indica a pesquisa. No trópico americano, a quantidade de zonas com vegetação alterada aumenta quanto mais longe elas estiverem dos limites dos territórios indígenas. Enquanto isso, uma indígena a determinada distância de uma floresta divisa os locais sob proteção formal, na África não se registrou diferença significativa. Existem locais destruídos tanto quanto dentro das terras indígenas.

Gardner explica que a riqueza de espécies de primatas é maior em terras indígenas e nas unidades de conservação ambiental, onde também há muitos macacos, principalmente nos trópicos americanos e na África indomalaia. "Mas isso não é o caso da África tropical, refletindo os impactos históricos do colonialismo. Os povos originais da África foram forçados a deixar suas terras tradicionais, ricas em biodiversidade, antes de essas áreas terem sido convertidas em unidades de conservação", explica o primatólogo norte-americano.

BUGIOS SOB ALTA PRESSÃO

No Brasil, a situação difere um pouco do que foi encontrado em outras partes do mundo. "A maioria das espécies que vivem na propriedade de distribuição das primatas (5%) inserida em terras de povos indígenas", informa o primatólogo Júlio César Bicca-Marques, da Pontifícia Universidade Católica de São Paulo (PUC-SP), e do Instituto Nacional de Pesquisas da Serra do Cipó (Inpe), autor do artigo. Embora o sistema nacional de unidades de conservação ambiental cubra outros

34% da área de ocorrência dos primatas, mais da metade dela (52%) está em regiões sujeitas a um maior impacto das atividades econômicas.

Segundo o estudo, o cenário pode mudar quando a distribuição de duas espécies de primatas no território brasileiro, o bugio-ruivo (*Aotus azarae*), comumente encontrado na Mata Atlântica, e o bugio-preto (*Aotus caraya*), que é mais comum no interior, muda. "Mais de 80% das áreas de ocorrência desses bugios estão fora de terras indígenas ou unidades de conservação. Isso significa que eles têm uma alta suscetibilidade a pressões humanas negativas", explica Bicca-Marques. "As populações também são vulneráveis à doença de febre amarela durante surtos silvestres da doença nas palmeiras fragmentadas onde elas vivem. Para elas, o exemplo do que ocorreu com os primatas griseus de amazônia, a perda de habitat com a avanço da agricultura, da urbanização e da construção de infraestrutura de transporte é crescente."

Autor: Wilson Spironello
Foto: Wilson Spironello, Instituto Nacional de Pesquisas da Amazonia (Inpa)

O primatólogo Wilson Spironello, do Instituto Nacional de Pesquisas da Amazônia (Inpa), concorda com a conclusão geral do trabalho, que destaca a importância das terras indígenas para a conservação da biodiversidade de macacos. No entanto, aponta que algumas peculiaridades também precisam ser levadas em conta, visto que certas espécies têm área de distribuição restrita. O que é preciso é que as populações de primatas variem dependendo do tamanho da terra indígena e do pressão de caça que esses animais sofrerem nas reservas", comenta Spironello, que não participou do estudo. "Os próprios indígenas consideram esses animais como fonte de recursos. É importante é analisar caso a caso."

Artigo científico
ESTRADA, A. et al. Global importance of Indigenous Peoples, their lands, and knowledge systems for saving the world's primates from extinction. *Science Advances* 10 ago. 2022.

PESQUISA FAPESP 320 | 87

Fonte: Pesquisa FAPESP, outubro de 2022, ano 23, n. 320, p. 56-57

Sob o título-rubrica²¹ “Ecologia” a reportagem é formada, na metade superior de sua primeira página, por uma foto de um menino com um macaco em seus braços. Essa imagem é acompanhada pela legenda “Criança da etnia Kalapalo segura macaco-prego de estimação no Parque Nacional do Xingu”. Foto e legenda apontam o espaço geográfico e cultural abordado no corpo da reportagem em que os macacos são protegidos e acolhidos pelo ser humano: as terras indígenas. O conjunto formado por fotos e títulos significa as terras indígenas como um espaço de acolhimento e harmonia entre seres humanos e animais.

Essa harmonia entre o macaco e os povos tradicionais representados na figura do menino da foto contrasta com o tema da destruição da biodiversidade do Brasil e do planeta, a qual é materializada no seguinte fragmento do corpo da reportagem:

²¹ Também denominado sobretítulo, antetítulo ou chapéu, o termo título-rubrica é utilizado aqui na acepção de Mouillaud et Tétu (1989, p. 118), enquanto “nomes sem determinante e desprovidos de predicados (...) Seu estatuto é comparável aos objetos de um catálogo ou às unidades de um repertório. (...) as categorias-rubricas do jornal são inscritas no meio cultural ao qual pertencem. Elas constituem, assim, uma fronteira entre o jornal e o mundo, um espaço de transição. Não é por acaso que os títulos-rubrica estão escritos nas laterais e no topo da página. Eles permitem a passagem do interior ao exterior do jornal.” [noms sans déterminant et dépourvus de prédicts (...) Leur statut est comparable à celui des objets d'un catalogue ou des unités d'un répertoire. (...) les catégories-rubriques du journal sont elles-mêmes inscrites dans l'environnement culturel auquel il appartient. Elles constituent ainsi une charnière entre le journal et le monde, un espace transitionnel. Il n'est pas sans signification que les titres-rubriques soient inscrits au bord et au sommet de la page. Ils permettent le passage de l'intérieur à l'extérieur du journal]. No nosso corpus, encontramos duas modalidades de rubricas: um primeiro conjunto constituído por títulos-rubrica que remetem a temas da realidade externa ao jornal (energia, agricultura, pecuária, biodiversidade, ecologia etc.), e um segundo composto por nomes de gêneros e/ou seções da revista (Notas, Memória, Entrevista etc.).

- (1) Mais de 90% das espécies cujo habitat não abrange terras indígenas estão ameaçadas de extinção. Entre as que podem ser encontradas nos territórios dos povos tradicionais, a proporção cai para 55%. “Quanto maior for a área de distribuição protegida, menor é o risco de as espécies de primatas serem classificadas como ameaçadas de extinção ou terem populações em declínio”, afirma Alejandro Estrada, do Instituto de Biologia da Universidade Nacional Autônoma do México, primeiro autor do trabalho. (Pesquisa FAPESP, out. 2022, ano 23, n. 320, p. 56. Grifos nossos)

Esse parágrafo da reportagem inicia com a síntese dos resultados de uma pesquisa da área de “biologia”, a qual é seguida pelo discurso direto de um dos autores do artigo científico citado. Conforme já abordamos acima, Valentín Volóchinov (2021[1929]) denomina esse modo de introdução do discurso alheio de “discurso direto preparado”, em que os principais temas do discurso direto são antecipados pelo contexto autoral e coloridos por suas entonações. Em primeiro plano, o contexto autoral da reportagem de popularização científica de *Pesquisa FAPESP* dá porcentagens de espécies ameaçadas fora de terras indígenas e as porcentagens menores em “territórios de povos tradicionais”. O discurso direto não traz propriamente informações novas, mas reforça aquelas já dadas no contexto autoral.

Identificamos aqui a ênfase valorativa que direciona a construção temática, composicional e estilística do enunciado. Essa ênfase está materializada em expressões como “por em risco”, “ameaçar”, “ameaçadas” e orientou, de modo recorrente, o conjunto dos enunciados sobre o meio-ambiente ou ecologia das 24 edições do *corpus* da pesquisa. A fim de comprovar a regularidade dessa ênfase valorativa, citaremos trechos de 10 enunciados de 10 diferentes edições do nosso *corpus*, que ficarão logo a seguir e não em anexo para facilitar sua visualização:

- (2) no corpo da Nota intitulada “Responsabilidade e diplomacia na ciência” (março 2021, ano 22, n. 301, p. 13): “Professor da Universidade de São Paulo (USP) e ex-diretor do Instituto Nacional de Pesquisas Espaciais (Inpe), Galvão recebeu neste ano o prêmio de Liberdade e Responsabilidade Científica, concedido a pesquisadores que põem em risco a carreira ou a segurança pessoal para resguardar a ciência em situações desafiadoras. Em meados de 2019, ele defendeu publicamente a qualidade e a confiabilidade de dados do Inpe que mostravam aumento expressivo no desmatamento da Amazônia dos ataques promovidos pelo presidente Jair Bolsonaro. Como consequência, Galvão foi exonerado.”
- (3) no corpo da reportagem sob o título-rubrica “Ambiente”, intitulada “Perdas ocultas” (abril de 2021, ano 22, no. 302, p. 65): “Desde o final da década de 1980, a área nativa de Mata Atlântica, um dos biomas mais ameaçados do país, tem se mantido relativamente estável”.
- (4) Reportagem sob o título-rubrica “Mudanças climáticas” intitulada “**Riscos de mais** desastres naturais (...) O aumento gradual do aquecimento gradual ao longo deste século deve intensificar progressivamente a incidência de chuvas extremas e elevar o risco de deslizamentos de terras e inundações bruscas nas regiões Sul e Sudeste e na faixa leste do Nordeste, onde ficam os maiores centros urbanos e se concentram mais de dois terços da população do Brasil.” (maio 2021, ano 22, n. 303, p. 64-65).

- (5) Reportagem sob o título-rubrica “Genética” intitulada “Refúgio ameaçado. Perda de vegetação nativa na Amazônia, onde vive quase metade das onças-pintadas do mundo, coloca em risco o maior reservatório genético da espécie” (agosto 2021, ano 22, n. 306, p. 60-61).
- (6) Reportagem sob o título-rubrica “Clima” intitulada “Cerrado ameaçado (...) O Cerrado, o segundo bioma brasileiro mais extenso e um dos mais ricos em diversidade de plantas e animais, encontra-se ameaçado” (novembro de 2021, ano 22, n. 309, p. 52-53).
- (7) reportagem de capa intitulada “Sob o risco da escassez (...) Além de reflexos na produção agrícola e no abastecimento de água nas cidades, a falta de chuvas colocou em risco a capacidade de geração de energia elétrica.” (dezembro 2021, ano 22, n. 310, p. 30-31).
- (8) fragmento de carta do leitor (janeiro de 2022, ano 23, n. 311, p. 6) “**Destrução do Cerrado.** “Ameaçado” é uma palavra inapropriada para o título da reportagem “Cerrado ameaçado” (edição 309). A destruição já é realidade e não mais um prenúncio, mesmo que ainda não seja total”.
- (9) no corpo da Nota intitulada “Mais árvores do que o imaginado” (março de 2022, ano 23, n. 313, p. 14): “‘Esses resultados destacam globalmente a vulnerabilidade da biodiversidade de florestas a mudanças antropogênicas, particularmente de uso da terra e do clima, porque a sobrevivência de espécies raras está desproporcionalmente ameaçada por essas pressões’, disse Peter Reich, um dos coordenadores do estudo, ao portal da Universidade de Michigan.”
- (10) no corpo da Nota intitulada “As florestas e a oferta de água nos EUA” (julho 2022, ano 23, n. 317, p. 15): “Os pesquisadores alertam que as queimadas, o desmatamento e as mudanças climáticas estão ameaçando a existência dessas florestas”.
- (11) no título da Nota “A ameaça da mineração em áreas indígenas” (novembro 2022, ano 23, n. 321, p. 10)

Os sentidos das expressões “põem em risco”, “elevar o risco”, “colocou em risco” bem como do verbo “ameaçar” e do substantivo “ameaça” apontam para dano ou prejuízo presentes a acontecer e podem despertar o medo ou temor no destinatário do enunciado. As perífrases verbais com o verbo auxiliar no presente e o principal no gerúndio ou no particípio (“estão ameaçando”, “está desproporcionalmente ameaçada”) indicam o “desenvolvimento do evento (aspecto cursivo)” (Neves, 2000, p. 63) e incidem sobre o destinatário que, movido pelo temor de destruição do meio-ambiente, poderia contribuir para a cessação do processo em curso. Desse modo, o verbo “ameaçar” no “aspecto verbal cursivo” e o substantivo dele derivado formam temas, que são determinados no contexto verbal e extraverbal dos enunciados de gêneros variados da divulgação científica (nota, reportagem, carta do leitor) e que, sem referência explícita ao locutor do enunciado, mobilizam o interlocutor-leitor a se posicionar diante da ameaça anunciada.

Desse conjunto, destacamos o excerto 2, que repetimos a seguir:

- (12) Professor da Universidade de São Paulo (USP) e ex-diretor do Instituto Nacional de Pesquisas Espaciais (Inpe), Galvão recebeu neste ano o prêmio de Liberdade e Responsabilidade Científica, concedido a pesquisadores que põem em risco a carreira ou a segurança pessoal para resguardar a ciência em situações desafiadoras. Em meados de 2019, ele defendeu publicamente a qualidade e a confiabilidade de dados do Inpe que mostravam aumento expressivo no desmatamento da Amazônia dos ataques promovidos pelo presidente Jair Bolsonaro. Como consequência, Galvão foi exonerado. (Pesquisa FAPESP, março 2021, ano 22, n. 301, p. 13, grifo nosso)

Nesse fragmento extraído do enunciado do gênero “Nota”, o jornalista anuncia o recebimento de um prêmio honorífico por um cientista e contextualiza esse recebimento com a informação de que, em 2019, ele sofreu “ataques promovidos pelo presidente Jair Bolsonaro”. O horizonte ideológico e a esfera política federal brasileiros orientam o estilo e a composição desse enunciado em que os cientistas estão “ameaçados” e “sob risco” por informarem sobre danos ao meio-ambiente. *Pesquisa FAPESP* evidencia aqui o embate político-ideológico entre, por um lado, o governo federal (2019-2022) e, por outro, parte da comunidade científica brasileira.

6 O ser humano ameaçado

Um segundo enunciado também presente na edição de outubro de 2022 (ano 23, n. 320, que reproduzimos mais abaixo, a fim de ficar próximo das análises do peritexto) é a reportagem da área da “Agricultura” intitulada “Agrotóxicos no trabalho.” Título auxiliar “Com o aumento do uso de pesticidas no país, trabalhadores rurais são mais expostos aos possíveis efeitos à saúde no curto e no longo prazo”.

Três aspectos da construção composicional colocam a reportagem em diálogo com o contexto político-ideológico mais amplo:

- 1 No corpo da reportagem, é abordado o embate de posições entre diferentes esferas por meio de formas de transmissão do discurso alheio;
- 2 O **peritexto** – títulos e fotografias – privilegiam a posição de um setor da esfera científica brasileira – pesquisadores do Departamento de Saúde Ambiental da Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP);
- 3 Um enunciado complementar inserido no meio da reportagem e intitulado “Polêmico projeto de lei” e com título auxiliar “Parlamento discute alterações nas regras de uso, comercialização e fiscalização de agrotóxicos” traz o contexto “defendido pela bancada ruralista”, que é um dos principais apoiadores do governo de Jair Bolsonaro.

O primeiro aspecto – o embate entre diferentes posições sobre os agrotóxicos – manifesta, por um lado, o contexto enquanto horizonte ideológico geral da sociedade brasileira contemporânea postulado por Pável Medviédev e, por outro, enquanto orientação autoral à palavra alheia, seja na mesma direção semântica, seja em direções semânticas distintas (polêmica, influenciando a palavra do autor de fora). Com base nessas duas dimensões do contexto, observamos que a reportagem abre com uma síntese dos resultados da investiga-

ção desenvolvida pelos pesquisadores do Departamento de Saúde Ambiental da Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP):

- (13) Dor de cabeça, taquicardia, fadiga, tontura, irritação das mucosas, vista embaçada e câimbra. Cerca de 90% dos participantes de um estudo feito com agricultores familiares de São José do Ubá, noroeste do Rio de Janeiro, em 2014 e 2015, apresentavam com frequência ao menos um desses sintomas, além de outros apontados, como decorrentes de intoxicação aguda por agrotóxico. Os pesquisadores do Departamento de Saúde Ambiental da Faculdade de Saúde Pública da Universidade de São Paulo (FSP-SP) que conduziram a investigação analisaram, além de sintomas de intoxicação aguda e crônica nessa comunidade, a saúde mental e respiratória dos 78 voluntários que participaram da pesquisa. Quase metade deles apresentava entre quatro e nove sintomas de intoxicação aguda e 25% tinham mais de quatro sintomas crônicos, como alteração do sono, irritabilidade, dificuldade de concentração e raciocínio. A região era a segunda maior produtora de tomate fluminense. “Os aplicadores de agrotóxicos, normalmente homens, e os ajudantes, em sua maioria mulheres, estão expostos a uma carga elevada dessas substâncias desde muito novos”, contou o idealizador do estudo Rafael Buralli, doutor em saúde pública pela USP. (Pesquisa FAPESP, out. 2022, ano 23, n. 320, p. 65)

O primeiro parágrafo da reportagem inicia com a síntese de uma pesquisa da área de saúde pública e é seguida pelo discurso direto do seu “idealizador”, ou seja, conforme vimos acima, o “discurso direto preparado” (Volóchinov, 2021[1929]). Em primeiro plano, o contexto autoral da reportagem de popularização científica de *Pesquisa FAPESP* enumera os diversos sintomas de intoxicação por agrotóxico e, em seguida, apresenta as porcentagens de agricultores participantes da pesquisa com todos ou alguns desses sintomas. Na sequência, o discurso autoral informa se tratar de uma pesquisa da Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP). Desse modo, o leitor é impactado pelos resultados dessa investigação, da qual o contexto autoral da revista se aproxima, por meio de ênfases valorativas de “resultados atestados”. Essa forma de transmissão do discurso alheio difere do modo como o contexto autoral da reportagem introduz, um pouco adiante, a voz de um representante do setor agrícola:

- (14) Reginaldo Minaré, diretor técnico-adjunto da Confederação da Agricultura e Pecuária do Brasil (CNA), reconhece que os pesticidas são produtos químicos perigosos e, por isso, “precisam ser manuseados com cuidado”. Mas ressalva que, considerando o volume aplicado no país, o trabalhador está até “bastante protegido”. “Temos a preocupação de preparar os agricultores e de fornecer assistência técnica para que façam uso adequado dessas substâncias, essenciais para a agricultura nacional”, diz. (Pesquisa FAPESP, out. 2022, ano 23, n. 320, p. 65)

Aqui o contexto autoral da reportagem coloca em primeiro plano a fonte das afirmações e, em seguida, utiliza a “modificação analítico verbal” do discurso indireto, em que “a personalidade aparece como uma *maneira* subjetiva (individual e típica), maneira de pensar e de falar, que inclui ainda a avaliação autoral dela” (Volóchinov, 2021[1929], p. 276). O discurso alheio é introduzido pelo verbo *dicendi* (ou de elocução ou introdutor de fala) “reconhece”,

que coloca a posição do setor agrícola na defensiva, para, em seguida, inserir entre aspas todas as expressões a respeito da segurança no manuseio dos agrotóxicos. As aspas apontam a maneira individual e típica de falar (Volóchinov, 2011[1929]) – a qual, nesse contexto, ganha o sentido de “na opinião dessa pessoa”, produzindo um distanciamento do contexto autoral em relação ao discurso alheio.

Na sequência, o discurso direto apresenta duas posições: em primeiro lugar, o autor defende que os agricultores têm orientação para o manuseio seguro de “agroquímicos”, substituindo o termo “tóxicos” da palavra composta “agrotóxicos” por “agroquímicos”, evitando, com isso, avaliações sociais negativas associadas a esse termo na língua portuguesa; e, em segundo lugar, ao afirmar que são “substâncias, essenciais para a agricultura nacional”, defende a necessidade dos “agroquímicos” para a viabilidade produtiva do setor. Esse argumento vai reaparecer em uma terceira voz da reportagem:

- (15) O engenheiro-agrônomo José Otávio Machado Menten, professor aposentado da Escola Superior de Agricultura Luiz de Queiroz (Esalq) da USP, em Piracicaba, ressalta que produtos químicos como fertilizantes e agrotóxicos são indispensáveis para o setor agrícola nacional. “Sem o uso de agroquímicos, o Brasil não seria a potência agrícola que é. Nossas projeções indicam que a produção nacional cairia pela metade se não usássemos defensivos no campo”, afirma. (*Pesquisa FAPESP*, out. 2022, ano 23, n. 320, p. 66 – grifos nossos)

Aqui novamente o contexto autoral traz, em primeiro plano, a identidade da fonte do discurso direto, um engenheiro-agrônomo e professor-pesquisador de uma prestigiada Faculdade de Agronomia. Em seguida, o contexto autoral novamente emprega o “discurso direto preparado”, só que, diferentemente do modo como transmite a voz do setor produtivo agrícola, faz uso da modificação analítico-objetual do discurso indireto. Essa forma prioriza os aspectos temáticos do discurso alheio e, embora Volóchinov (2018[1929]) afirme que há uma distância clara entre o discurso alheio e o autoral, a não utilização das aspas para afirmar que os agrotóxicos são “indispensáveis”, aproxima a informação de um fato, que, em seguida, é reforçado pelo discurso direto do engenheiro-agrônomo.

A reportagem é construída por meio do diálogo entre três vozes-posições sociais: a da esfera científica da saúde pública, a da esfera do setor produtivo do agronegócio e a da esfera científica da engenharia agrícola. A posição inicial na reportagem bem como o maior espaço concedido à voz da pesquisa em saúde pública parecem apontar que o jornalismo científico de *Pesquisa FAPESP* assume uma ênfase valorativa que se aproxima da voz-posição social da esfera da saúde pública, ao mesmo tempo que reconhece as necessidades do setor produtivo do agronegócio brasileiro.

O segundo aspecto da reportagem que manifesta sua relação com o contexto são os componentes do seu peritexto: título-rubrica, título, título auxiliar e fotografia no início da reportagem.

Imagen 2 – Agrotóxicos no trabalho



Fonte: Pesquisa FAPESP, outubro de 2022, ano 23, n. 320, p. 64-65

O título-rubrica “AGRICULTURA” conecta o enunciado da reportagem com seu contexto espacial e cultural em dois níveis: primeiramente, sua localização na extremidade superior esquerda o situa em um espaço de transição entre a revista e o mundo; em segundo lugar, o título-rubrica é formado por um nome ou substantivo comum desprovido de determinantes e predicados que “só se resolvem na função de referênciação (...) na própria instância da construção do enunciado” (Neves, 2000, p. 73), associando o tema da reportagem ao contexto espacial e cultural brasileiro: o campo (em oposição à cidade), o da produção de alimentos e um dos principais setores de exportação e de produção de divisas para o Brasil, o chamado “agronegócio”.

O título “AGROTÓXICOS NO TRABALHO” é composto por um substantivo no plural – cuja significação única no dicionário é “produto de origem química ou biológica usado na prevenção ou extermínio de pragas e doenças das culturas agrícolas (fungicidas herbicidas, inseticidas, pesticidas); agroquímico, defensivo agrícola” (Houaiss, 2009, p. 72) – acompanhado por um substantivo no singular precedido pela preposição articulada a um artigo definido, que delimitam simultaneamente uma atividade e o local de sua realização. A disposição espacial desse título é feita em duas linhas: a primeira com o substantivo “AGROTÓXICOS” apresenta o tema geral da reportagem; e a segunda com o advérbio de lugar “NO TRABALHO” que, em conjunto com a fotografia de abertura da reportagem onde está introduzida, significa ao mesmo tempo a atividade produtiva e seu local de realização.

No primeiro plano da fotografia, um jovem vestido de camiseta, bermuda e boné caminha em uma plantação ao mesmo tempo que carrega um recipiente às costas. Desse recipiente sai uma mangueira que é segurada pelo jovem com uma das mãos e que emite uma fumaça branca. Na margem inferior direita da foto e em letras bem pequenas, aparece a descrição “Agricultor pulveriza inseticida em lavoura de município da Região Metropolitana de São Paulo”.

AGRICULTURA

Com o aumento do uso de pesticidas no país, trabalhadores rurais são os mais expostos aos possíveis efeitos à saúde no curto e no longo prazo

Frances Jones e Yuri Vasconcelos

Dora, coquelicite, taquicardia, fadiga, náusea, vômito, dor de cabeça, dor de estômago e cefaleia. Cerca de 60% dos participantes de um estudo feito com agricultores familiares de São José do Uba, no norte do Rio de Janeiro, em 2014 e 2015, apresentavam com frequência ao menos dez desses sintomas, além de outros apontados como decorrentes de intoxicação aguda por agrotóxicos. Os pesquisadores do Departamento de Saúde Ambiental da Escola de Medicina da Universidade de São Paulo (USP-USP) que conduziram a investigação analisaram, além de sintomas de intoxicação aguda e crônica nessa comunidade, a saúde mental e respiratória dos 78 voluntários que participaram da pesquisa. Quase 70% delas apresentavam de quatro a nove sintomas de intoxicação aguda e 25% tinham mais de quatro sintomas crônicos, como alteração da sonolência, irritabilidade, dificuldade de concentração e raciocínio. A agrotóxica era a segunda maior preocupação da comunidade.

“Os aplicadores de agrotóxicos, normalmente homens, e os ajudantes, em sua maioria mulheres, estão expostos a uma carga elevada dessas substâncias desde muito novos e o idealizador da estatística, Bernardo Uebel, doutor em estatística pela USP. A pesquisa resultou em quatro artigos científicos e foi o tema da tese de doutorado de Buralli, que hoje é diretor técnico da Coordenação-geral de Vigilância à Saúde do Trabalhador (Cvst) do Ministério da Saúde. As perguntas sobre quais agrotóxicos usavam, os agricultores citaram 49 diferentes pesticidas de 31 grupos químicos – entre eles alguns proibidos para o uso de tomate, o que já havia sido banido no Brasil.

À intoxicação aguda por agrotóxicos de uso agrícola, que em 2017 provocou 61 das mortes registradas no país pelo Sistema Nacional de Informações de Mortes por Causa Biológica (Sinc), é segundo estendendo os assuntos, aponta a parte mais visível dos efeitos à saúde que podem ser causados às pessoas diretamente expostas aos pesticidas. Em 2017, último ano que se tem registro, 31 mil casos foram contabilizados, 2.548 casos de intoxicação aguda.

A intoxicação aguda por agrotóxicos se manifesta por meio de sintomas e sintomas clínicos de efeito nocivo resultantes da interação do produto com os indivíduos responsáveis por seu manejo ou uso e/ou exposição. Ela se apresenta de forma súbita, por vezes com a exposição aguda ao agente químico. Já a intoxicação crônica resulta da exposição regular, durante meses ou anos, a agrotóxicos.

O Instituto da Agricultura, Pecuária e Abastecimento (Mapa) informou, por meio de sua Assessoria Especial de Comunicação Social, que não é possível atribuir o número total de mortes por intoxicação em 2017 ao maior uso de agrotóxicos, o que teria ocorrido devido pelas próprias tabelas do Sincrônico, dos 61 óbitos registrados de tentativas de suicídio, dois foram reportados como acidente e os demais ocorreram em “circunstâncias ignoradas”, ou seja, não esclarecidas na morte.

Reginaldo Minaré, diretor técnico adjunto da Confederação da Agricultura e Pecuária do Brasil (CNA), reconhece que os pesticidas são produtos químicos perigosos e, por isso, “precisam ser usados com cuidado”. Minaré afirma que, considerando o volume aplicado no país, o trabalhador está “bastante protegido”.

“Temos a preocupação de preparar os agricultores e de fornecer a competência técnica para que fiquem adequadamente informados, essenciais para a agricultura nacional”, diz. O Serviço Nacional de Aprendizagem Rural (Senar) capacitou mais de 200 mil trabalhadores rurais para o uso de agrotóxicos na última década, segundo Minaré. No Brasil, 100 mil pessoas trabalham em estabelecimentos agropecuários, conforme o Censo Agro 2017, do Instituto Brasileiro de Geografia e Estatística (IBGE). O diretor da CNA defende que é preciso ampliar a conscientização sobre o risco correto dos efeitos da exposição ao agrotóxico desrespeito dos trabalhadores. Estamos caminhando para um patamar interessante de segurança.”

PESQUISA FAPESP 220 | 65

Em gramáticas da língua portuguesa, a posição anteposta ao substantivo “marca a interveniência de uma avaliação subjetiva do falante na qualificação efetuada” (Neves, 2000, p. 203). A partir de uma visão da linguagem como atividade discursiva que, assim como Rastier, considera a língua como um sistema “incessantemente modificado pelo uso e trabalhado por dinâmicas históricas (Rastier, 2022[1998], p.168), o adjetivo “polêmico” inscreve a ênfase valorativa dos autores da reportagem da revista *Pesquisa FAPESP*—Frances Jones e Yuri Vasconcelos, respectivamente editor de tecnologia e colaborador – a qual destaca a disputa entre diferentes setores da sociedade brasileira.

O sintagma nominal “projeto de lei” que sucede o adjetivo “polêmico” bem como o substantivo “Parlamento” na posição de sujeito ativo do título-auxiliar do enunciado inscrevem a disputa na esfera da política legislativa federal brasileira. Com isso, essa esfera funciona como contexto político motivador da seleção do tema científico da reportagem maior – a divulgação dos resultados de pesquisa sobre saúde física e mental de agricultores realizada pela Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP). Em outros termos, embora do ponto de vista da construção composicional o debate legislativo tenha sido introduzido como um aspecto contextual secundário, é o contexto político da esfera legislativa que pode ter motivado a escolha da pesquisa e orientado sua organização estilística e composicional por meio de aspectos verbais e não verbais (fotos, diagramação etc.) que materializam a ênfase valorativa dos autores do dossiê e dos editores do periódico de divulgação científica *Pesquisa FAPESP*.

O texto dessa reportagem dentro da reportagem se organiza para apresentar as duas posições em disputa:

Quadro 1— Projeto de lei (PL) no. 1.459/2022²²

“Chamado de PL do Veneno pelos críticos”	“Defendido pela bancada ruralista”
Para a Anvisa ²³ , a quem cabe avaliar os aspectos toxicológicos, de risco à exposição ocupacional e dietética dos agrotóxicos, o PL no. 1.459 “enfraquece a regulação de produtos agrotóxicos no país, especialmente a avaliação do impacto desses produtos para a saúde humana dos consumidores de alimentos”. (grifos nossos)	Reginaldo Minaré, diretor técnico-adjunto da Confederação da Agricultura e Pecuária do Brasil (CNA), <u>avalia que o projeto é oportuno</u> . “A insatisfação do setor agrícola com o sistema de registro, absolutamente moroso, vem de 2005, pelo menos”, destaca. Segundo ele, o registro de um novo agrotóxico no país leva por volta de oito anos. (grifos nossos)

Fonte: *Pesquisa FAPESP*, n. 320, out. 2022, p. 68

Sob a forma do discurso direto preparado, o contexto autoral de *Pesquisa FAPESP* informa a posição social dos locutores de ambos os lados da polêmica. No caso do discurso da Anvisa, a avaliação geral do projeto, “enfraquece a regulação”, aparece no discurso direto alheio. Quanto ao representante da bancada ruralista (conjunto de parlamentares que defendem interesses do agronegócio brasileiro no Congresso Nacional), sua avaliação está no contexto autoral “avalia que o projeto é oportuno” e sua justificação – “a insatisfação do setor

²² O “PL dos Agrotóxicos” foi aprovado pela Comissão de Agricultura (CRA) no dia 19/12/2022 e enviado ao plenário do Senado, onde se encontrava à época em processo de tramitação e ainda não tinha sido votado. Informações disponíveis em: <https://www12.senado.leg.br/noticias/videos/2022/12/pl-dos-agrotóxicos-segue-para-o-plenário>. Acesso em: 01/05/2023

²³ Agência Nacional de Vigilância Sanitária.

agrícola devido à morosidade do processo de registro” – é *dada pelo discurso direto*. Desse modo, a ênfase valorativa presente no contexto autoral do corpo da reportagem de *Pesquisa FAPESP* aproxima-se, reconhece e valoriza a voz do setor agrícola, muito embora o título apresente o tema como sujeito a discussão, disputa, polêmica.

7 Conclusões

A noção de contexto desempenha um papel-chave na definição de uma epistemologia enunciativa, textual, discursiva e sociológica da linguagem: por um lado, sua atualidade é atestada pelas formulações de Dominique Maingueneau (1993), François Rastier (2002[1998]) e Teun van Dijk (2012[2007]); por outro, sua complexidade nas ciências da linguagem se expressa na posição de Mikhail Bakhtin (2017[1970–71], p. 44) de que o “contexto é potencialmente inacabável”, o que pudemos verificar acima pelos vários aspectos do contexto presentes no método sociológico de M. Bakhtin, P. Medviédev e V. Volóchinov.

O contexto adquire variadas nuances nas diversas obras e cronologias de Bakhtin e seu Círculo. Para este trabalho, concentramo-nos nas acepções presentes nos textos da segunda metade dos anos 1920, momento da constituição do método sociológico por Mikhail Bakhtin, Pável Medviédev e Valentin Volóchinov. O contexto orientou a análise da divulgação científica de *Pesquisa FAPESP* em duas dimensões:

- 1 O discurso ou contexto verbal autoral, em que a revista transmite, assimila, interpreta e avalia o discurso alheio. Nessa acepção, constatamos um predomínio da modificação “modelo discurso direto preparado” (Volóchinov, 2021[1929]), por meio do qual o contexto autoral introduziu posições em conflito sobre o tema do meio ambiente, aproximando-se de algumas, distanciando-se de outras, e às vezes sugerindo a necessidade de uma conciliação entre elas;
- 2 A parte extraverbal do enunciado é constituída, no caso de nosso *corpus* de análise, sobretudo pelas esferas ideológicas envolvidas (principalmente científica, política e jornalística) e pelo horizonte ideológico geral da sociedade brasileira entre 2021 e 2022 dominado por um governo federal que opôs desenvolvimento econômico e ecologia e atacou discursos científicos que denunciavam, entre outros, o aquecimento climático e defendiam um desenvolvimento sustentável com base na preservação dos biomas brasileiros. Direcionadas por avaliações sociais mais amplas, as ênfases valorativas dos autores dos enunciados de popularização científica conectam a parte verbal (mediante a escolha do conteúdo temático, dos recursos estilísticos e da construção composicional) e a não verbal (fotos, imagens, diagramação) com a parte extraverbal dos enunciados (contexto ideológico e político).

Entre janeiro de 2021 e dezembro de 2022, as edições de *Pesquisa FAPESP* se concentraram em temas de preservação do meio-ambiente orientados por ênfases valorativas de ameaça, risco aos biomas brasileiros e à saúde da população.

Por meio de uma análise pormenorizada de um número, colocado em relação com enunciados de diferentes gêneros que trataram do meio-ambiente entre 2020 e 2021, totalizando 24 edições, procuramos conectar a singularidade dos enunciados analisados a um con-

texto mais amplo a fim de chegarmos a avaliações sociais (posições ideológicas constitutivas do horizonte ideológico amplo de uma época e de uma sociedade) norteadoras do contexto político e ideológico brasileiro e a ênfases valorativas dos autores dos enunciados.

Referências

- BAKHTIN, M. M. *Os gêneros do discurso*. Trad. Paulo Bezerra. São Paulo: Editora 34, 2016[1953-54].
- BAKHTIN, M. Fragmentos dos anos 1970–1971. In: BAKHTIN, M. *Notas sobre literatura, cultura e ciências humanas*. Org., trad., posf. e notas P. Bezerra. São Paulo: Editora 34, 2017[1970–71]. p. 21–56.
- BAKHTIN, M. *Problemas da poética de Dostoiévski*. Org. trad., posfácio e notas P. Bezerra. 5e ed. Rio de Janeiro: Forense Universitária, 2010[1963].
- BAKHTIN, M. *Questões de estilística no ensino da língua*. Trad. S. Grillo e E. V. Américo. São Paulo: Ed. 34, 2013.
- BAKHTIN, M. *Problemas da obra Dostoiévski*. Trad., notas e glossário. S. Grillo et E. V. Américo. São Paulo: Editora 34, 2022[1929].
- BOUQUET, S. De l'hexagramme cognitiviste à une sémiotique de l'interprétation. In: RASTIER, F.; BOUQUET, S. *Une introduction aux sciences de la culture*. Paris: PUF, 2002. p. 11-35.
- CHARAUDEAU, P.; MAINGUENEAU, D. *Dictionnaire d'analyse du discours*. Paris: Seuil, 2002.
- DEBRAY, R. *Cours de médiologie générale*. Paris: Gallimard, 1991.
- GENETTE, G. *Seuils*. Paris: Seuil, 1987
- GREIMÀS, A. J.; COURTÈS, J. *Dicionário de semiótica*. Trad. A. D. Lima et al. São Paulo: Contexto, 2008[1993].
- GRILLO, S. V. de C. *A produção do real em gêneros do jornal impresso*. São Paulo: Humanitas/Fapesp, 2004.
- GRILLO, S. V. de C. Esfera e campo. In: BRAIT, B. *Bakhtin: outros conceitos-chave*. São Paulo: Contexto, 2006. p. 133-160.
- GRILLO, S. V. de C. *Divulgação científica: linguagens, esferas e gêneros*. 2013. 333 f. Tese (Livre-Docência) – Faculdade de Filosofia, Letras e Ciências Humanas, São Paulo, 2013.
- GRILLO, S. V. C. A noção de ‘contexto’ na obra de Mikhail Bakhtin e do Círculo. ALFA: revista de linguística, v. 67, p. 1-28, 2023. DOI: <https://doi.org/10.1590/1981-5794-e17813>
- GRILLO, S. V. de C.; AMÉRICO, E. V. Registros de Valentin Volóchinov nos arquivos do ILIAZV. In: VOLÓCHINOV, V. *A palavra na vida e a palavra na poesia: ensaios, artigos, resenhas e poemas*. Org., trad., notas S. Grillo e E. V. Américo. São Paulo: Editora 34, 2019. p. 7-56.
- HOUAIS, A., VILLAR, M. DE S. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2009.
- MAINGUENEAU, D. *Le contexte de l'œuvre littéraire. Énonciation, écrivain, société*. Paris: Dunod, 1993.
- MEDVIEDEV, P. *O método formal nos estudos literários*. Introdução crítica a uma poética sociológica. Trad. e notas E. V. Américo e S. Grillo. São Paulo: Contexto, 2012[1928].
- MOUILAUD, M.; TETU, J.-P. *Le journal quotidien*. Lyon: Presses Universitaires de Lyon, 1989.

- NEVES, M. H. M. *Gramática de usos do português*. São Paulo: UNESP, 2000.
- RASTIER, F. Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage. *Languages*, Paris, v. 32, n. 129, p. 97-11, 1998. Disponível em: https://www.persee.fr/doc/lge_0458-726x_1998_num_32_129_2149. Acesso em 20/08/22.
- SPERBER, D.; WILSON, D. *Pertinence, communication & cognition*. 2. ed. Oxford Royaume-Uni/ Cambridge États-Unis: Blackwell, 1995[1986].
- SCHLEIERMACHER, F.D.E. *Herméneutique*. Pour une logique du discours individuel. TC Berner, Alençon (Orne): CERF/PUL, 1987[1809-1810].
- SPINELLI, D. A dialética texto e contexto em Senhora, de José de Alencar ou Considerações sobre Literatura e Sociedade, de Antônio Cândido. *Kalíope*, São Paulo, v. 04, n. 7, p. 29-47, 2008.
- VANDENDORPE, C. Contextes, compréhension et littérature. *RSSI*, Ottawa, v.1, n. 11, p. 9-25, 1991.
- VAN DIJK, T. A. *Discurso e contexto*: uma abordagem sociocognitiva. Trad. R. Ilari. *São Paulo, Contexto*, 2012[2007].
- VOLÓCHINOV, V. N. *Marxismo e filosofia da linguagem*. Problemas fundamentais do método sociológico na ciência da linguagem. Tradução, notas e glossário S Grillo e E. V. Américo. 3. ed. São Paulo: Editora 34, 2021[1929].
- VOLÓCHINOV, V. A palavra na vida e a palavra na poesia: para uma poética sociológica. In: VOLÓCHINOV, V. *A palavra na vida e a palavra na poesia*: ensaios, artigos, resenhas e poemas. Org., tradução, ensaio introdutório e notas S. Grillo e E. V. Américo. São Paulo: Editora 34, 2019[1926]. p. 109–146.
- VOLÓCHINOV, V. O que é a linguagem/língua? In: VOLÓCHINOV, V. *A palavra na vida e a palavra na poesia*: ensaios, artigos, resenhas e poemas. Org., tradução, ensaio introdutório e notas S. Grillo e E. V. Américo. São Paulo: Editora 34, 2019a[1930]. p. 234–265.
- VOLÓCHINOV, V. Estilística do discurso literário II: A construção do enunciado. In: VOLÓCHINOV, V. *A palavra na vida e a palavra na poesia*: ensaios, artigos, resenhas e poemas. Org., tradução, ensaio introdutório e notas S. Grillo e E. V. Américo. São Paulo: Editora 34, 2019b[1930]. p. 266–305.
- VOLÓCHINOV, V. Estilística do discurso literário III: A palavra e sua função social. In: VOLÓCHINOV, V. *A palavra na vida e a palavra na poesia*: ensaios, artigos, resenhas e poemas. Org., tradução, ensaio introdutório e notas S. Grillo e E. V. Américo. São Paulo: Editora 34, 2019c[1930]. p. 306–336

Parâmetros vocais e prosódicos na fala de influenciadores digitais transgênero da mídia social brasileira

Vocal and Prosodic Parameters on Transgender Digital Influencers' Speech on Brazilian Social Media

Geovana Soncin

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
geovana.soncin@unesp.br
<https://orcid.org/0000-0003-4903-1919>

Eryne Alves Bafum

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
eryne.a.bafum@unesp.br
<https://orcid.org/0009-0007-5059-2808>

Gabriela Aparecida Rodrigues Gonçalves

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
rodrigues.goncalves@unesp.br
<https://orcid.org/0000-0002-9291-586X>

Giovanna Caroline Borges

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
g.borges@unesp.br
<https://orcid.org/0009-0001-7128-6515>

Karoline Araujo dos Santos

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
karoline.araujo@unesp.br
<https://orcid.org/0000-0002-6756-1289>

Eliana Maria Gradim Fabbrion

Universidade Estadual Paulista (UNESP)
Marília | São Paulo | SP | BR
eliana.fabbrion@unesp.br
<https://orcid.org/000-0001-5197-0347>

Resumo: Analisaram-se parâmetros vocais e prosódicos na fala de influenciadores digitais brasileiros transgêneros, por meio de avaliação perceptivo-auditiva e análise acústica para identificar semelhanças e diferenças. Também foram comparados os padrões acústicos de fala entre homens e mulheres transgênero (trans) e cisgênero (cis). Amostras de fala analisadas foram constituídas por áudios de doze vídeos selecionados, baseados em temas e visualizações, publicados no Instagram. Na avaliação perceptivo-auditiva, em avaliação cega, os juízes avaliaram como julgavam o gênero do falante e identificaram fronteiras prosódicas e proeminências usadas para fins expressivos. Na análise acústica, mensurou-se a frequência fundamental para caracterização vocal do falante, a variação de f_0 e a produção de pausas como parâmetros indicativos de fronteiras prosódicas e, ainda, variação de f_0 e de intensidade para marcação de proeminências. Observou-se que a fala dos sujeitos trans, seja por sua expressão vocal, seja por suas características prosódicas, não é avaliada exclusivamente por meio da recuperação de padrões acústicos. Na comparação dos parâmetros acústicos mensurados na fala de sujeitos trans e cis, observou-se que a fala de sujeitos trans é dialeticamente constituída, pois existem pontos de aproximação e distanciamento entre a caracterização de aspectos vocais e prosódicos na fala desses sujeitos em relação à fala do gênero declarado.

Palavras-chave: representações identitárias; transgênero; avaliação auditiva; acústica; prosódia.



Abstract: Voice and prosody in speech of transgender Brazilian digital influencers were analyzed, using perceptual assessment and acoustic analyses to identify similarities and differences. Acoustic speech patterns between transgender (trans) and cisgender (cis) men and women were also compared. Speech samples analyzed consisted of audios from twelve selected videos published on Instagram. In a blinded perceptual assessment, judges evaluated how they judged the speaker's gender and identified prosodic boundaries and prominences used for expressive purposes. In the acoustical analysis, the voice was measured by the fundamental frequency; prosodic boundaries were measured by f_0 variation of pause production; also, prominence marking was measured by f_0 variation and intensity. Results show that transgender subjects' speech, either by their vocal expression or by their prosodic characteristics, is not perceived exclusively through the recovery of acoustical patterns. Acoustic parameters measured in speech of transgender subjects in comparison with speech of cisgender subjects showed that transgender speech is dialectically constituted, since there are points of approximation and distancing between the characterization of vocal and prosodic aspects of their speech and the characterization of these aspects in speech of the declared gender.

Keywords: identity representations; transgender; auditory assessment; acoustical analysis; prosody.

1 Introdução

Este artigo apresenta análises de parâmetros prosódicos de influenciadores digitais brasileiros que se declaram como sujeitos trans. Por um lado, o artigo procura caracterizar prosodicamente a fala desses sujeitos, tanto do ponto de vista da avaliação perceptivo-auditiva, ou seja, sobre como essa fala é julgada auditivamente por sujeitos outros, quanto do ponto de vista acústico, a fim de verificar quais seriam pontos de congruência e/ou incongruência entre a caracterização acústica e o julgamento atribuído à sua fala. Por outro lado, o artigo discute questões relacionadas ao papel do gênero, enquanto socialmente construído e linguisticamente manifestado, na subjetividade das pessoas trans.

No interior do movimento político e social identificado pela sigla LGBTQIA+¹, o termo “transgênero” não se refere a uma orientação sexual, mas é utilizado de forma ampla para descrever pessoas que se identificam de maneira diversa em relação ao seu sexo biológico, incluindo transgêneros (homem ou mulher), travestis (termo que designa a performance de uma identidade feminina) ou, ainda, pessoa não-binária, que se comprehende além da divisão “homem e mulher” (Ministério dos Direitos Humanos, 2018; Reis, 2018). Sujeitos transgêneros², portanto, nascem com um sexo biológico que não corresponde ao gênero com o qual se identificam. Por exemplo, uma mulher trans é alguém que nasceu com o sexo biológico masculino, mas se reconhece como mulher. Da mesma forma, um homem trans é designado como pertencente ao sexo biológico feminino ao nascer, mas se identifica como homem.

Assim, comprehende-se que pessoas transgênero, tanto homens quanto mulheres vivem identidades de gênero que diferem das expectativas culturais associadas ao sexo biológico. A identidade de gênero, nesse sentido, está relacionada à forma como o sujeito se percebe e se identifica internamente, seja como feminino ou masculino. Trata-se, portanto, de uma experiência subjetiva que define com qual gênero a pessoa se alinha e como se declara (Governo do estado de São Paulo, 2015).

Para que se possa no presente artigo discutir questões linguísticas relacionadas aos sujeitos transgênero, o conceito de gênero adquire papel central no presente estudo, uma vez que o termo “transgênero” é dele derivado. De acordo com Matos (2008)³, o gênero como um conceito surgiu na década de 70 e disseminou-se na década seguinte no contexto científico. Sua proposição teve como objetivo dissociá-lo do conceito de *sexo*, criando, assim, a distinção entre gênero e sexo. Nessa diferenciação, enquanto *sexo* é “uma categoria analítica marcada pela biologia e por uma abordagem essencializante da natureza ancorada no biológico” (Matos, 2008, p. 336), *gênero* é uma “dimensão que enfatiza traços de construção histórica, social e sobretudo política que implica análise relacional” (Matos, 2008, p. 336).

Embora a distinção entre gênero e sexo tenha sido um ganho no campo científico, a proposta de um sistema de classificação social que se ancora no conceito de gênero, conforme apresenta Matos (2008), é privilegiadamente acionada de forma binária (masculino *versus* feminino) e raramente em formato tripartite. Com efeito, a binariedade se transforma em norma presente na sociedade, no seio da qual constrói-se discursivamente o que é “normal” e “anormal” no que diz respeito ao gênero. Ou seja, quando um sujeito cuja identificação de gênero não é compatível com o modelo binário, como é o caso dos sujeitos transgênero, alvos

¹ O Manual de Comunicação LGBTI+, organizado pela Aliança Nacional LGBTI+, elenca os significados das letras que compõem a sigla LGBTQIA+, na qual a letra T remete aos sujeitos transgênero. Aos interessados nas demais populações que integram a sigla, sugerimos conferir o manual, indicado na lista de referências do presente texto como Reis (2018).

² O termo transgênero inclui o termo transexual embora dele se diferencie, uma vez que o último envolve além da mudança social de gênero transição somática por tratamento hormonal e cirurgia genital. Em contrapartida, transgênero é um termo mais amplo que indica a maneira diversa com que uma pessoa identifica seu gênero em relação ao sexo biológico sem a necessidade de tratamento hormonal ou cirúrgico, pois, muitas vezes esses procedimentos não são realizados nem mesmo indicados, já que a identidade de gênero é uma questão social e não patológica. Por essa razão, o termo transexual não é recomendado embora ainda alguns usos equivocados persistam (cf. a propósito: Dias, 2021). Por esse entendimento, nesse trabalho, adotamos “transgênero”.

³ O trabalho de Matos (2008) é oriundo do campo das ciências sociais. Em seu texto, a autora reporta como o conceito de gênero é construído no interior do movimento feminista e reflete como esse conceito promove uma ruptura nos modos tradicionais de pensamento sobre a estrutura social.

do presente estudo, rompe-se com o sistema predominante socialmente imposto, o que pode ser - equivocadamente - interpretado como *anormal*.

Os juízos de valor *normal* e *anormal* são frequentes em discursos correntes no Brasil no que diz respeito à população trans, haja vista exemplos amplamente divulgados pela imprensa nacional quanto a preconceitos sofridos por esse grupo social⁴. Em direção contrária, vários dados apontam que a população brasileira é diversa e heterogênea em sua constituição não apenas no que diz respeito a povos e etnias que estão na base de sua formação social, mas também no que tange à identidade de gênero e de orientação sexual, uma vez que no Brasil a população LGBT equivale a 12% (Spizzirri *et al.*, 2022) da sua população e tem grande expressividade no mundo. Dados divulgados pelo Google, por exemplo, apontam que o Brasil foi, em 2022, o terceiro país no ranking mundial com maior índice de busca pelo termo LGBTQIA+ e, ainda, que nos últimos cinco anos houve aumento de 90% na busca por esse termo (Alves, 2022). Dados como esses demonstram a necessidade de compreensão dessa população, que mostra cada vez mais a sua existência e suas necessidades.

Observa-se, no entanto, que apesar de sua heterogeneidade, a sociedade brasileira é normativamente binária no que diz respeito ao gênero, ou seja, impõe, aos sujeitos, a cis-generidade binária como “normal”. Tem-se, nessa sociedade, portanto, um sistema em que a “norma” quanto à identidade de gênero, como masculino ou feminino, deveria ser correlata à identificação dada pelo sexo biológico, como homem ou mulher, respectivamente. Dessa relação normativa, é interesse do presente artigo chamar atenção para o fato de que esse sistema dissemina os seus valores simbólicos na distinção binária feminino/masculino em várias materialidades, como roupas, brinquedos, modos de andar e de se comportar e, também, em modos de falar (Drumond, 2009).

Ressalta-se particularmente neste artigo que a fala é um dos meios pelos quais um sujeito pode ser identificado quanto ao seu gênero pelas suas características de natureza indicial, ou seja, a fala dá indícios sobre as características do aparelho fonoarticulatório do falante bem como sobre o seu estilo de fala na comunicação oral, permitindo classificá-lo quanto ao seu gênero por essas características (Barbosa; Madureira, 2015; Lima; Constantini, 2017). Nessa perspectiva, compõe a fala o que se reconhece como voz, mas, longe de ser esse um conceito definido pela fisiologia, Barros Filho (2005) o comprehende a partir do olhar da interação social com o outro mediada pela linguagem. Desse modo, para o autor, voz não se trata de um veículo físico disponível para a comunicação. Ao contrário, segundo Barros Filho (2005), voz é parte integrante da manifestação verbal porque faz parte da significação do enunciado, tendo, assim, efeito valorativo para a interpretação deste pelo outro. Nas palavras do autor, “voz não é veículo porque é mensagem” e, para tanto, a voz é constituída socialmente e não herdada (Barros Filho, 2005, p. 34).

É sobre esse aspecto que o presente artigo se debruça ao desenvolver uma análise que teve como objetivos (i) comparar o julgamento perceptivo-auditivo atribuído à fala de sujeitos transgênero influenciadores digitais brasileiros com os parâmetros acústicos de natureza vocal e prosódica que caracterizam a fala desse sujeitos; (ii) comparar os parâmetros acústi-

⁴ Em 2022, houve registro de pelo menos 131 casos de homicídios de pessoas trans, abrangendo 130 travestis e mulheres trans e 1 homem trans (Benevides, 2023; conferir em <https://antrabrasil.files.wordpress.com/2023/01/dossieantraz2023.pdf>); destaca-se também o caso de transfobia em relação à deputada federal Erika Hilton durante votação na Câmara sobre casamento homoafetivo no Brasil, que foi amplamente noticiado (conferir em: <https://queer.ig.com.br/2023-09-19/erika-hilton-sofre-transfobia-em-votacao-de-pl-contra-casamento-lgbt.html>

cos que caracterizam vocal e prosodicamente a fala de sujeitos transgênero em relação à fala de sujeitos cisgêneros, especificamente compara-se homens trans em relação a homem cis e mulheres trans em relação à mulher cis.

Ao atender o objetivo (i), o artigo irá comparar se o modo como a fala dos sujeitos trans é julgada pela avaliação perceptivo-auditiva corresponde a como ela se caracteriza acusticamente. Por sua vez, ao atender o objetivo (ii), o artigo irá responder se a fala dos sujeitos transgênero se aproxima ou se distancia das características acústicas da fala do gênero declarado, ou seja, serão adotados critérios que permitam identificar se a fala de homens trans se aproxima ou não da fala do gênero masculino e se a fala de mulheres trans se aproxima ou não da fala do gênero feminino, considerando o padrão cisgênero como referência.

2 Aspectos teóricos

2.1 A fala e suas características prosódicas: funções indiciais e linguísticas

A forma como os sujeitos trans se expressam verbalmente é essencial para afirmar sua identidade de gênero. Além disso, as características acústicas da fala, como a frequência fundamental, desempenham um papel importante na identificação do gênero de quem fala. Portanto, compreender a relação entre as características atribuídas à fala dos sujeitos trans por meio de um julgamento perceptivo-auditivo e as características acústicas é fundamental para entender a complexa intersecção entre identidade e expressão vocal.

Segundo Cagliari (1992), o estilo de fala de um falante envolve vários elementos prosódicos, sendo a entonação especialmente ligada à estrutura sintática. Além disso, a entonação desempenha um papel crucial na expressão das atitudes do falante. Por outro lado, a velocidade da fala e o volume da voz também são aspectos importantes (Cagliari, 1992) que indicam estados afetivos do falante e dão pistas sobre as situações de fala em que os enunciados são produzidos.

Um dos parâmetros acústicos já discutidos largamente na literatura para a identificação do gênero do falante é a frequência fundamental (f_0), tanto para pessoas cisgênero, quanto para transgênero. Em revisão de literatura, envolvendo estudos de diversas nacionalidades, na faixa etária entre 18 e 50 anos, foram observados resultados de f_0 entre 190 e 251 Hz para mulheres e entre 112 e 141 Hz para homens (Spazzapan, *et al.*, 2019). Em falantes jovens adultos do Português Brasileiro (18 a 49 anos), essa medida variou de 198 Hz a 223 Hz para mulheres e de 121 Hz a 129 Hz para homens, na emissão sustentada da vogal “a” (Spazzapan *et al.*, 2022). Entretanto, o intervalo de 150 Hz a 185 Hz é reconhecido como uma faixa neutra (*neutral range*) ou ambígua para a identificação de gênero (*gender ambiguous*), conforme discutido por Hardy *et al.* (2018). Destaca-se que os termos “faixa neutra” ou “gênero ambíguo” são traduções dos termos originalmente usados no inglês pelos autores, ainda que, de nossa parte, se saiba que esses termos, nos estudos da linguagem, possam ser problematizados.

Em sujeitos transgêneros, estudos realizados no Brasil identificaram f_0 de 172,4 Hz em mulheres trans (MT), com base na avaliação da emissão sustentada da vogal /ε/ e na contagem numérica de um a 20 (Schmidt *et al.*, 2018).

Em pessoas transgênero, estudos apresentaram valores de f_0 em amostras de fala de mulheres trans (MT) que variaram entre 118 Hz e 233 Hz (Mcneill *et al.*, 2007); em termos de média, diferentes valores foram identificados por estudos distintos, com é o caso das médias iguais a 163 Hz (Van Borsel; De Pot; De Cuypere, 2007); 172,40 Hz (Schmidt *et al.*, 2018); 128,31 Hz (Vieira, 2018); e 159,046 Hz (Villas-Bôas *et al.*, 2021). Estudos que avaliaram a fala de homens trans (HT) são menos numerosos, porém Silva *et al* (2021) apresentaram valores de f_0 que variaram entre 122,51 Hz e 174,26 Hz nesse grupo.

No que diz respeito à comparação entre sujeitos trans e cis, o estudo de Santos e Antunes (2020) apresentou como resultado que valores de f_0 de mulheres trans foram muito próximos aos valores de mulheres cis, todas elas locutoras de vídeos produzidos como conteúdo digital e disponíveis no Youtube. Com esse resultado, o estudo defende que a semelhança identificada entre as vozes de mulheres trans e cis se explica principalmente pela construção social da voz (Barros Filho, 2005) e muito menos por aspectos de natureza fisiológica, como o caso da categoria 'sexo', pois as vozes podem ser construídas também de forma performática de modo a subverter o pré-construto de que sexo, gênero e sexualidade se relacionam de forma consonante. Assim, o estudo conclui que categorias vocais e prosódicas não são binárias nem heteronormativas, mas sim complexas por serem também afetadas por questões sociais.

Em relação à identificação do falante pelo julgamento perceptivo-auditivo da fala de homens e mulheres transgênero, ouvintes não treinados identificaram os falantes como vozes femininas em 64% dos julgamentos, enquanto a identificação das vozes desses sujeitos como vozes masculinas e indefinidas foi equivalente a 25,8% e 9,7% respectivamente (Schmidt *et al.*, 2018). Hardy *et al.* (2018) descreveram que juízes avaliaram vozes masculinas com valor médio de f_0 de 122,93 Hz, vozes femininas com 194,60 Hz e, ainda, que o valor de 173,94 Hz foi assinalado como gênero ambíguo. Os autores relatam ainda que obtiveram 85% de acerto nos julgamentos realizados.

Refletindo ainda sobre marcos prosódicos que diferenciam a fala entre homens e mulheres, diversos estudos investigaram essas diferenças em padrões de duração da vogal (Ericsdotter, Ericsson, 2001; Hillenbrand *et al.*, 1995; Johnson, Martin, 2001;) e verificaram que mulheres produzem vogais com maior duração que homens. Especificamente, foi relatado um padrão consistente na língua sueca, em que falantes do sexo feminino produziram maiores diferenças duracionais entre *tokens* tônicos e átonos de vogais nas mesmas palavras monossilábicas (Ericsdotter, Ericsson, 2001).

No que diz respeito a outros parâmetros prosódicos, pesquisadores apontaram que a entoação é um marcador de fala e que pode distinguir grupos de diferentes identidades de gênero (Papeleu *et al.*, 2023). Entretanto, estudos relacionados ao gênero envolvendo a marcação de fronteira prosódica e a marcação de proeminência – conforme é realizado no presente trabalho – são raros.

É consensual na literatura que a função demarcativa e a função de marcação de proeminência são duas funções primordiais desempenhadas pela prosódia na organização da fala para fins comunicativos (D'imperio *et al*, 2015; Hirst; Di Cristo, 1998; Ladd, 1996; Lehiste, 1979; Levelt, 1989; Pijper; Sanderman, 1994; Swerts; Geluykens, 1994, entre outros). A função demarcativa é responsável por marcar o limite de unidades por meio do estabelecimento de

fronteiras prosódicas, indicando, assim, o quanto tais unidades estão unidas ou separadas na cadeia da fala. Por sua vez, a função de marcação de proeminência é responsável por distinguir informações importantes daquelas que são menos importantes, criando assim pontos proeminentes na fala (Terken; Hermes, 2000). Em outras palavras, a demarcação de fronteiras prosódicas indica as transições entre unidades linguísticas, atuando na segmentação da fala e favorecendo a interpretação semântica dos enunciados falados, enquanto a proeminência destaca unidades específicas no interior desses enunciados.

Estudos mostraram que as fronteiras prosódicas, marcadas foneticamente por pausas, alongamento pré-fronteira e mudanças na frequência fundamental e no espectro sonoro, ajudam na identificação de unidades linguísticas, como palavras e frases (Cutler; Norris, 1988; Serra, 2009, Soncin, 2018). Essas fronteiras fornecem pistas acústicas e contextuais que facilitam a compreensão do discurso e a organização da informação (Breen *et al.*, 2010).

Por sua vez, a proeminência contribui para a expressividade do sujeito falante frente ao seu dizer, uma vez que permite a focalização prosódica em diferentes pontos do enunciado. Diversos estudos mostram que, em diferentes línguas, a proeminência é alcançada por meio de características acústicas como aumento de duração e de intensidade, maior magnitude de frequência fundamental e padrões melódicos específicos (Astésano, *et al.*, 2004; Gussenhoven, 2006, Terken; Hermes, 2000, entre outros). Para o Português, essa descrição é apresentada por Moraes (2009); Barbosa e Madureira (2015); Carpes (2019), Santos *et al.* (2023), entre outros. A marcação de proeminência permite aos falantes enfatizar palavras ou expressões específicas, direcionando a atenção do ouvinte para informações importantes. Como efeito, a marcação de proeminência desempenha um papel essencial na expressão efetiva de ênfase retórica, emoção e intenção comunicativa (Patterson, 2019).

Apesar da relevância da marcação de fronteiras prosódicas para a organização da fala e da marcação de proeminência para efeitos expressivos e semânticos, não foram encontrados trabalhos que tenham abordado esses aspectos na fala da população transgênero. Não obstante, no que diz respeito à identificação de falantes pela medida de f_0 , outro fator de análise considera no presente trabalho, os achados da literatura são inconclusivos e, ainda, não necessariamente relacionam análises acústica e perceptivo-auditiva, o que torna desejável a realização de novos estudos para melhor se compreender a fala dessa população.

Nesse sentido, o presente estudo se insere nos estudos de caracterização da fala da população trans e, nesse contexto, discute o desdobramento dessa caracterização para a construção de uma representação identitária dos sujeitos trans. Por essa razão, a seguir, apresenta-se a perspectiva teórica a partir da qual o sujeito transgênero é abordado neste trabalho considerando a relação eu/outro.

2.2 O sujeito transgênero à luz de uma concepção dialógica de linguagem

Nos estudos linguísticos contemporâneos, especialmente considerando a inserção da pesquisa linguística no âmbito dos Objetivos do Desenvolvimento Sustentável da Agenda 2030 da Organização das Nações Unidas (ONU), a emergência de questões políticas e pautas sociais são inevitáveis. “Igualdade de gênero” é um dos dezessete temas da Agenda 2030 e sobre ele a discussão acerca do fenômeno linguístico, considerando a relação entre linguagem, língua e sujeito, tem muito a contribuir, haja vista a luz que essa discussão lança a questões relacio-

nadas às relações intersubjetivas instanciadas e marcadas pela linguagem/língua e, ainda, ao efeito dessas relações às representações sociais e identitárias.

É sabido que, em diferentes níveis de análise e quadros teóricos, aspectos linguísticos são sistematicamente submetidos a metodologias investigativas para demonstrar a sua condição de constituinte social. Por exemplo, conforme apresenta Pinto (2012), variações sintáticas e fonológicas são estudadas sob a ótica da significação social que assumem para falantes e/ou ouvintes, os relatos de mulheres são interpretados no que indiciam de suas autoimagens e das imagens que o ponto de vista masculino tem delas e, ainda, com outros objetivos estudos mostram que “o ensino de línguas é analisado à luz dos processos coloniais e de globalização” (Pinto, 2012, p. 75).

Neste artigo, o percurso investigativo se encaixa no perfil, identificado por Pinto (2012), de uma pesquisa que, sob um procedimento metodológico delineado, busca descrever como aspectos linguísticos de natureza prosódica (e, portanto, de natureza fonético-fonológica) caracterizam a fala de sujeitos transgênero produzida em contexto de mídias sociais e, ainda, busca indicar como essa fala é julgada no que diz respeito à sua constituição social. Interessa, pois, ao presente trabalho a constituição social do sujeito trans, em sua relação mútua com a constituição linguística, uma vez que, de acordo com Matos (2008) o gênero é uma dimensão que se define por traços de natureza social, histórica e política (cf. Introdução).

Tal proposta investigativa, embora formal do ponto de vista da análise linguística, tendo em vista a caracterização acústica das amostras de fala dos grupos de homens e mulheres trans (o detalhamento a esse respeito é dado na seção destinada ao método), ancora-se num quadro teórico que concebe a linguagem e o sujeito como dialógicos, conforme os fundamentos de Bakhtin (1981, 2003), com especial filiação à releitura desses fundamentos apresentada por Ribeiro e Sobral (2021) para a discussão da constituição das representações identitárias.

Ribeiro e Sobral (2021) abordam a constituição das representações identitárias à luz da Teoria Dialógica do Discurso (Bakhtin; Volóchinov, 1976, 2004; Bakhtin, 1981, 2003, entre outros), em interface com a Teoria das Representações Sociais (Marková, 2006; Moscovici, 1961, entre outros). Da abordagem de Ribeiro e Sobral (2021), três pilares são destacados aqui: a linguagem como constitutivamente dialógica; a constituição dos sujeitos como necessariamente marcada pela relação *eu-outro*; e o vazio na relação *eu-outro* como espaço simbólico para a emergência das representações identitárias.

Considerar a linguagem como constitutivamente dialógica implica concebê-la não como estática e abstratamente sistemática, mas, ao contrário, como fluxo significativo que se estabelece pela relação entre enunciados, na medida em que qualquer enunciado é resposta a outro enunciado, que, por sua vez, abrirá possibilidades para outros enunciados-resposta. Ocorre, porém, que os enunciados são ditos por sujeitos, os quais são afetados pelo social e pelo histórico. Assim, a dinâmica dialógica da linguagem ocorre porque o diálogo se realiza no âmbito social e, para tanto, os sujeitos são chamados como atores sociais a produzir discurso, ou seja, a linguagem é dialógica porque os sujeitos que nela e por ela se constituem são sujeitos responsivos e, portanto, são sujeitos que respondem de forma valorativa ao(s) enunciados dos outro(s).

Tal responsividade valorativa é real, pois os sujeitos se constituem pela *relação eu-outro*. Por essa razão, sua realidade nunca é individual. É somente a partir da identificação de uma relação deslocada em relação ao outro que o sujeito define a si mesmo como um eu e, a partir desse lugar socialmente construído, expressa suas atitudes valorativas frente ao mundo. Conforme palavras de Ribeiro e Sobral (2021),

O sujeito é membro de uma coletividade e, mais especificamente, de um grupo e de um segmento específico dessa coletividade e, sendo influenciado pelo outro em sua constituição como sujeito, e influenciando a constituição dos outros com quem interage, seus enunciados não descrevem simplesmente o mundo, mas o descrevem de acordo com as maneiras como seu grupo e seu segmento social veem o mundo. [...] A maneira como é sua existência social e histórica constitui os modos como ele [sujeito] vê o mundo e, assim, os modos como valora esse mundo. E esses modos se manifestam no discurso. (Ribeiro; Sobral, 2021, p. 6)

Considerando a assunção da relação eu-outro como constitutiva do sujeito, Ribeiro e Sobral (2021) dissertam sobre as funções identitárias. Para os autores, com base no aspecto dialógico da linguagem e no aspecto sempre relacional – e nunca isolado – do sujeito, não existiria uma identidade *a priori*, ou seja, uma identidade pré-existente, mas sim *representações identitárias* que se marcam no discurso, “como fruto de lugares sociais assumidos diante do deslocamento do sujeito-membro em relação a seu grupo de pertença” (Ribeiro; Sobral, 2021).

Particularmente, os autores chamam atenção para o fato de que as representações identitárias emergiriam em novas formas, ou configurando novas ordens, quando o deslocamento do sujeito em relação a seu grupo (o outro) encontra o vazio como espaço simbólico. Ou seja, na medida em que não existiriam identidades fixas e visadas pelos sujeitos, representações identitárias seriam tecidas por relações eu/outro que se caracterizariam pela incompletude, pois as lacunas identificadas entre o eu e o outro fariam emergir novas representações simbólicas sobre si. Nesse caso, assumindo a relação entre o dialógico e a relação simbólica entre homem e mundo, os autores propõem que o vazio “é o lugar das possíveis percepções a serem elaboradas através da emergência de esboços diferentes uns dos outros, pois também a nossa singularidade é cunhada sob a égide da alteridade” (Ribeiro; Sobral, 2021).

Na esteira dessas considerações, neste trabalho os sujeitos transgêneros são vistos como constitutivamente marcados por um vazio simbólico entre o eu e o outro na realidade social na medida em que esses sujeitos recusam como seu gênero aquele que seria o gênero correlato ao seu sexo de nascimento numa relação supostamente biunívoca. A ruptura que faz emergir uma nova realidade social é simbólica, uma vez que o sujeito identifica uma lacuna entre o que é o gênero do outro e o que se esperaria, histórica e socialmente, que fosse o seu (numa correlação direta com o sexo), e, assim, a transforma em lugar de subjetividade ao manifestar uma representação identitária outra: aquela do sujeito trans. Sujeito esse que responde dialogicamente aos outros sobre a relação sexo/gênero a partir de um lugar valorativo construído que se marca na linguagem pela relação de alteridade: no caso de homens trans, o “eu sou” equivale à não identificação com o corpo biológico que lhe foi dado, pois o gênero que se vê socialmente associado ao corpo feminino lhe falta, fazendo emergir o gênero masculino como identitário; no caso de mulheres trans, o processo seria o mesmo, mas em ordem inversa, pois o “eu sou” equivale à não identificação com o corpo e o gênero masculino, pois o que socialmente se representa como masculino lhe é lacunar, provocando como efeito a identificação com a representação social que se tem do feminino. O preenchimento simbólico das lacunas identificadas na relação eu-outro permite, assim, que a relação sujeito-mundo-linguagem se reconfigure nesse lugar simbólico distinto de representação identitária de si.

À luz desse modo de conceber os sujeitos trans, entendemos que a linguagem é crucial para refletirmos sobre a questão da identidade e do sujeito, e, assim fazemos nossas as afirmações de Ribeiro e Sobral (2021):

Vemos então que as representações identitárias também evidenciam o pressuposto dialógico, visto que sempre há uma sombra do outro (pares ou grupos) para a acentuação de quem se é enquanto membro ou o que se é enquanto grupo. O outro balizará os limites de quem eu sou na dialética do movimento. O vazio, diante do quadro interacional, é condição para a responsividade, é quando o sujeito joga a bola no espaço para o outro rebater. Novos valores são (re)colocados, conferindo à trama uma nova configuração, novos vazios, para novas respostas. (Ribeiro; Sobral, 2021, p. 19)

Nessa conjuntura, podemos, pois, considerar que, discursivamente, as novas respostas à sociedade dadas pelos sujeitos trans consistem na afirmação de sua condição identitária como trans, ou seja, ao afirmarem “eu sou homem trans” ou “eu sou mulher trans”, esses sujeitos se inserem na língua e atualizam novos sentidos sociais e históricos, configurando uma nova realidade de linguagem e de sociedade. Essa nova configuração ganha proporções maiores quando, em movimento de grupo, esses sujeitos utilizam as mídias sociais para marcar sua presença como sujeitos, haja vista a repercussão e o alcance de público que as mídias sociais, enquanto suporte, atingem.

3 Material e Método

3.1 Constituição da amostra

Para a constituição da amostra analisada, foi realizado levantamento de vídeos postados na plataforma Instagram no período de janeiro a fevereiro de 2023 por sujeitos transgênero e cisgênero, especificadamente: homens trans (HT) e mulheres trans (MT), homem cis (HC) e mulher cis (MC)

Os critérios de seleção adotados foram: (i) vídeos postados que tratassesem de dois temas, a saber, “Visibilidade Trans” e “Retificação de nome e gênero”, por serem temas que proporcionaram variadas postagens da população trans no Instagram; (ii) quantidade de visualizações informada pela plataforma de mídia Instagram, sendo assim, dos vídeos disponíveis na plataforma sobre os temas definidos, foram selecionados os vídeos com maior número de visualizações (o intervalo de visualizações foi de 540 mil a 20 mil); (iii) idade dos sujeitos entre 18 e 40 anos. A partir desses critérios, foram selecionados doze vídeos: dez vídeos de sujeitos trans autodeclarados, dos quais cinco foram produzidos por homens trans e cinco por mulheres trans, e dois vídeos de sujeitos cis (um produzido por homem cis e um por mulher cis). A escassez de vídeos de homens e mulheres cis discutindo os temas selecionados foi uma limitação na constituição da amostra, já que não foram encontrados variados vídeos de sujeitos cis tratando desses temas.

Após selecionados, os vídeos foram baixados em formato mp4 e, posteriormente, os áudios de cada vídeo foram extraídos, salvos em arquivos individuais, convertidos em formato wav, inspecionados e analisados pelo *software* PRAAT (Boersma; Weenink, 2005). Foram também retirados sons de vinhetas para a realização da análise acústica. Destaca-se que a extração dos áudios foi realizada para que a análise fosse realizada sem informações de natureza visual e se centrasse nas informações de natureza auditiva. Assim, os áudios obtidos

foram submetidos à análise acústica. Anteriormente, porém, em etapa prévia definida como procedimento metodológico, os áudios foram também submetidos à avaliação perceptivo-auditiva de juízes conforme detalhamento a seguir.

3.2 Avaliação perceptivo-auditiva

A avaliação perceptivo-auditiva é amplamente utilizada em pesquisas na área da saúde, mais especificamente na Fonoaudiologia e em estudos envolvendo a avaliação da voz. Neste tipo de procedimento metodológico, juízes experientes ou leigos fazem julgamentos de um determinado parâmetro vocal e/ou de fala. É explicitamente considerada uma avaliação subjetiva pois é dependente da percepção do avaliador e, por isso, neste procedimento, pode-se utilizar comparações inter e/ou intra-avaliador. Estudos com a população LGBTQIA+ envolvendo a qualidade vocal e/ou a fala têm utilizado tal procedimento metodológico (Canal, *et al.*, 2024; Leung, Oates, Chanb, 2018).

Neste estudo, a avaliação perceptivo-auditiva, realizada por meio de juízes, foi adotada como etapa metodológica para identificar, por um lado, como o gênero dos sujeitos trans influenciadores, sujeitos falantes das amostras de fala sob análise, seria julgado auditivamente e, por outro, para identificar pontos julgados pelos juízes como referentes a fronteiras prosódicas e proeminências relevantes para a expressão de fala desses sujeitos para os fins enunciativos na plataforma digital. Primeiramente, um grupo de juízas leigas, que se designavam como cisgênero, avaliou com percebiam o gênero dos participantes. Em seguida, um grupo de juízas especialistas, que também se designavam como cisgênero, identificou fronteiras prosódicas e proeminências nas falas dos participantes. Essa etapa foi adotada, tendo em vista um dos objetivos do trabalho que busca comparar os resultados da análise acústica das amostras de fala desses sujeitos com a avaliação perceptivo-auditiva.

Os procedimentos adotados para a realização da avaliação perceptivo-auditiva são descritos abaixo.

3.2.1 Julgamento do gênero do falante

Para essa avaliação na qual os juízes julgaram o gênero do falante, os 12 áudios que compõem a amostra analisada (relembre-se sendo cinco de MT, cinco de HT, um de homem cis e um de mulher cis) foram editados no *software PRAAT* (Boersma; Weenink, 2005) a fim de considerar apenas 10 segundos de cada gravação de forma a controlar o conteúdo da fala para que não houvesse a declaração verbal explícita do gênero. Na seleção desses 10 segundos, deu-se preferência aos momentos em que não ocorria identificação pessoal e/ou uso de itens lexicais que remetesse ao gênero declarado. Adotou-se 10 segundos por ser suficiente para avaliação quanto ao gênero, enquanto tempo superior ou com integralidade dos áudios não permitiria o sigilo quanto ao gênero declarado devido às escolhas lexicais, sigilo esse necessário para essa avaliação. A edição controlada dessa forma foi importante para manter a avaliação de forma cega quanto ao conteúdo e, ainda assim, com tempo suficiente para a avaliação dos juízes.

Depois de editados, os áudios foram organizados em arquivos sem identificação e enviados para o julgamento perceptivo-auditivo sobre o gênero do falante para três juízes leigos, ou seja, sem conhecimento sobre o tema do trabalho e sem conhecimento técnico especializado sobre análise de voz e fala, sobretudo prosódia. Para tanto, os juízes foram

recrutados no primeiro ano de um curso de Fonoaudiologia, do interior do estado de São Paulo, pois, nessa etapa de formação, os alunos não cursaram disciplinas teóricas sobre análise acústica, voz e prosódia. Optar por juízes leigos é uma decisão metodológica que tenta assegurar que o julgamento esteja mais próximo à percepção do senso comum, menos afeita a vieses técnicos e mais alinhada com as experiências cotidianas da sociedade, ainda que esse efeito perceptual seja atravessado por questões não controladas pelo estudo. Esta abordagem está em consonância com as práticas padrão em pesquisas realizadas com juízes, mantendo a relevância e aplicabilidade dos resultados.

Para a realização da avaliação pelos juízes, os áudios foram apresentados por um questionário elaborado no *Google Forms*. Como tarefa, os juízes deveriam ouvir os áudios e classificar cada voz como feminina, masculina ou indefinida, ou seja, como uma voz que poderia ser considerada tanto como feminina como masculina.

3.2.2 Julgamento de fronteiras prosódicas e marcação de proeminência

A avaliação perceptivo-auditiva nesta fase do estudo foi utilizada para que pudessem ser levantados os usos de proeminências e de marcação de fronteiras para então seguir para análise acústica desses pontos identificados perceptualmente. Nessa fase da avaliação perceptivo-auditiva, os doze áudios que compõem a amostra foram julgados por um novo grupo de juízes, dessa vez com conhecimento técnico especializado. Compuseram o grupo três avaliadores com formação em fonoaudiologia, conhecimento e experiência na análise de parâmetros prosódicos, especialmente no que diz respeito à marcação de fronteira e de proeminência. Esses juízes foram ainda tutorados por uma quarta avaliadora, linguista e especialista em análise prosódica. A escolha de especialistas visou um julgamento mais técnico e aprofundado para a avaliação de aspectos prosódicos, evitando, assim, a necessidade de introduzir conceitos de fronteira e proeminência prosódica aos juízes caso fossem leigos e, por consequência, procurando manter a confiabilidade metodológica do estudo em relação à percepção desses aspectos prosódicos. Nessa etapa da avaliação, foi importante trabalhar com uma amostra de fala maior de cada sujeito para a identificação das fronteiras e proeminências, diferentemente das amostras usadas no julgamento do gênero do falante, que foram editadas numa duração mais curta para não ser explicitado o gênero declarado por itens lexicais.

Dos 12 áudios, 11 deles foram editados para que cada um tivesse a mesma duração de 39 segundos. Como eram áudios com duração maior, o tempo de 39 segundos foi selecionado como índice de padronização temporal. Um dos áudios, porém, permaneceu com 21 segundos, sua duração máxima original. Para a organização do material a ser submetido a julgamento pelos juízes, realizou-se a transcrição ortográfica de cada áudio, sem uso de sinais de pontuação e maiúsculas.

Os áudios e as transcrições de fala foram organizados, sem identificação, em arquivos alocados no *Google Drive* e compartilhados com os juízes. Cada juiz recebeu a instrução de ouvir cada gravação, quantas vezes fossem necessárias, e anotar no texto transscrito as fronteiras prosódicas que julgavam identificar auditivamente e as palavras que, sob a sua percepção, foram produzidas com proeminência prosódica para fins expressivos. Nas raras situações em que não houve concordância de pelo menos dois juízes, a quarta juíza, linguista, experiente na tarefa exigida, realizou a avaliação. Assim, para a identificação das fronteiras e das palavras proeminentes, foram consideradas, dessa forma, as anotações coincidentes por, pelo menos, duas avaliações.

3.3 Análise acústica

A análise acústica teve por objetivo mensurar os parâmetros acústicos que definem a voz e caracterizam fronteiras prosódicas e a marcação de proeminência nos áudios de cada sujeito da amostra. A seguir, detalham-se quais parâmetros acústicos foram considerados para caracterizar esses aspectos e como foram mensurados.

3.3.1 Análise da Frequência Fundamental

A extração do valor médio de f_0 da fala de cada participante foi realizada utilizando o Software PRAAT (Boersma; Weenink, 2005). Para essa análise, foram analisados 15 segundos de fala a partir do início da fala do participante. Uma inspeção visual anterior à mensuração de f_0 foi necessária, por meio da análise espectrográfica e auditiva, para remover ruídos (como sons de respiração, interrupções na fala e risadas), vinhetas e qualquer tratamento na voz com efeitos sonoros.

3.3.2 Caracterização acústica das fronteiras prosódicas

Foi realizada análise acústica das fronteiras que foram identificadas nas amostras de cada sujeito pelos juízes. Esse procedimento se justifica a fim de que se possa cotejar se o que foi avaliado no julgamento perceptivo-auditivo, em etapa metodológica anterior, correspondeu ou não a mudanças nos parâmetros acústicos na fala dos sujeitos, uma vez que a relação entre percepção e manifestação acústica não necessariamente é idêntica, ou seja, perceber a fala não se confunde com a recuperação de um padrão acústico (cf. a esse respeito: Liberman; Mattingly, 1985; Fowler, 1996; Goldstein; Fowler, 2003). Para análise acústica das fronteiras, considerou-se como parâmetros a variação de f_0 e a presença de pausa silenciosa, extraíndo-se, também, sua duração.

3.3.2.1 Análise da variação de frequência fundamental nas fronteiras

As fronteiras identificadas na fala apresentada em cada áudio da amostra, foram analisadas pela extração dos valores mínimo e máximo de f_0 . Para a extração desses valores, a sílaba tônica da palavra anterior a cada fronteira foi identificada e a partir dela, foram extraídas manualmente as referidas medidas pelo Software PRAAT com o uso da ferramenta *Get Pitch* (Boersma; Weenink, 2005).

3.3.2.2 Análise da produção e duração de pausas nas fronteiras

Ainda para a caracterização das fronteiras, foi realizada a análise da presença e da duração de pausa silenciosa no local onde as fronteiras foram identificadas pelos juízes especialistas. Esse procedimento visou compreender a relação entre as fronteiras identificadas e a presença de pausa silenciosa. Embora existam outros tipos de pausa na literatura prosódica (Merlo; Barbosa, 2012; Zellner, 1994), como a pausa hesitativa e a pausa preenchida, a pausa silenciosa desempenha função de segmentação da fala em unidades sintático-semânticas e/ou prosódicas e ocorre em fronteiras de constituintes prosódicos (Barbosa, 2006), como o enun-

ciado fonológico e a frase entoacional (Nespor; Vogel, 1986). Diferentemente, as pausas hesitativas são um subtipo de marcas linguísticas que caracterizam o fenômeno da hesitação e, do ponto de vista da organização prosódica, tendem a ocorrer em posições não coincidentes com fronteiras desses constituintes (Nascimento; Chacon, 2006). Considerando que a análise acústica buscou caracterizar as fronteiras prosódicas identificadas pelos juízes, analisa-se, pois, a pausa silenciosa, pois trata-se do tipo de pausa esperado para os limites prosódicos identificados no julgamento realizado.

As amostras de fala de cada sujeito foram analisadas, inicialmente, para verificar a presença ou ausência de pausa nas fronteiras identificadas. A presença ou não de pausas foi analisada por meio de observação visual do espectrograma do sinal acústico no *software PRAAT* (Boersma; Weenink, 2005). No espectrograma foi identificada a presença ou não de silêncio, pela análise do traçado. Para tanto, foi considerada, em cada áudio, a produção de fala imediatamente após a palavra que foi assinalada como fronteira no julgamento perceptivo-auditivo. A partir dessa avaliação, foram contabilizadas as fronteiras identificadas que foram marcadas acusticamente por pausa silenciosa. As pausas identificadas foram submetidas à extração manual da duração, considerando o final da palavra que antecedeu a fronteira identificada até o início da palavra posterior à fronteira. Os valores obtidos foram então organizados em número de pausas e tempo de duração de pausa em segundos.

3.3.3 Caracterização acústica das proeminências

Após a conclusão da análise das fronteiras, o estudo avançou com foco nas proeminências que os juízes identificaram. Cada uma dessas proeminências passou por uma análise acústica detalhada, concentrando-se na sílaba tônica das palavras envolvidas, nas quais mensurou-se a variação de frequência fundamental e a variação de intensidade.

3.3.3.1 Análise da variação de frequência fundamental das proeminências

As proeminências identificadas em cada amostra foram examinadas acusticamente de maneira isolada por meio da extração manual da f_0 mínima e máxima da sílaba tônica das palavras identificadas como proeminentes. Para tanto, utilizou-se a ferramenta *Get Pitch* do *software PRAAT* (Boersma; Weenink, 2005). A sílaba tônica foi identificada manualmente pelo traçado spectrográfico e as medidas mínima e máxima de f_0 foram anotadas.

3.3.3.2 Análise da variação de intensidade das proeminências

Da mesma forma, as palavras que foram identificadas como proeminentes na avaliação perceptivo-auditiva, foram analisadas acusticamente para a extração manual do parâmetro de intensidade. Para tanto, fez-se a extração da intensidade mínima e máxima da sílaba tônica das palavras proeminentes com o uso da ferramenta *Get Pitch* do *software PRAAT* (Boersma; Weenink, 2005); para a análise de variação de intensidade, considerou-se a diferença do valor máximo em relação ao valor mínimo.

3.4 Forma de análise dos dados

Os dados foram analisados de forma descritiva, utilizando abordagens quantitativa e qualitativa. Não foi feita análise estatística inferencial devido ao número restrito de sujeitos cujas amostras de fala foram analisadas. Nesse sentido, os resultados apresentados são descritivos e não tem intenção preditiva ou generalizante, haja vista os objetivos do estudo de natureza mais qualitativa e, ainda, da especificidade dos dados analisados: amostras de fala de sujeitos trans influenciadores e, portanto, não oriundos da população trans em geral.

Foi feita inicialmente uma organização dos aspectos observados na avaliação perceptivo-auditiva em etapa metodológica anterior à análise acústica para que se pudesse, por fim, comparar os resultados da análise acústica com o julgamento atribuído à fala dos sujeitos trans e cis pelos juízes, dados os objetivos do presente artigo.

Para organizar os resultados do julgamento perceptivo-auditivo do gênero dos sujeitos, foi calculada a porcentagem de respostas dos juízes. Para o julgamento das fronteiras prosódicas, foram calculados os índices de fronteira, que correspondem à relação entre o número de fronteiras identificadas pelos juízes e o número de palavras produzidas na amostra analisada de cada sujeito. A média e o desvio padrão (DP) foram calculados para cada grupo de participantes trans. De forma semelhante, a análise das proeminências foi realizada através do cálculo do índice de proeminências, considerando a relação entre o número de palavras proeminentes identificadas e o total de palavras produzidas. A média e o DP foram também calculados para os grupos de participantes trans.

Em relação à análise acústica, os valores de f_o de cada participante foram apresentados junto com a média e o DP de cada grupo, HT e MT, e os valores de HC e MC foram apresentados em números absolutos para cada participante. A análise acústica das fronteiras prosódicas, especificamente os resultados da variação de f_o , foi organizada em tabelas e gráficos a partir dos valores médios de cada sujeito. Os resultados referentes à frequência de pausa nos pontos identificados como fronteiras prosódicas foram apresentados em porcentagem e a duração das pausas, em segundos, foi apresentada pela média para cada participante. No que se refere às proeminências, os resultados da variação de f_o e da variação de intensidade foram exibidos em tabelas e gráficos por meio dos valores médios de cada sujeito.

As comparações entre os resultados do julgamento perceptivo-auditivo e as medidas acústicas são apresentadas de forma qualitativa.

Foram assumidas duas hipóteses: 1) Considerando que a avaliação perceptivo-auditiva de aspectos vocais e expressivos da fala é complexa pois envolve aspectos linguísticos mais amplos para além da recuperação de características acústicas, assume-se que o julgamento sobre o gênero na expressão vocal de sujeitos transgêneros não seria correspondente necessariamente os parâmetros acústicos relacionados ao gênero declarado; no que diz respeito às características prosódicas, embora existam possíveis correlações entre parâmetros acústicos e a avaliação perceptivo-auditiva, assume-se ainda que haveria discrepâncias entre o que é identificado auditivamente e o que se mostra acusticamente na fala da população trans que indicariam uma complexa interação entre a identidade de gênero e a expressão verbal; 2) Considerando que os sujeitos trans se definem pela relação dialética que estabelecem com outros sujeitos – especialmente com os sujeitos do gênero cuja identidade é congruente com o sexo de nascimento –, assume-se que a fala de sujeitos transgêneros apresentaria diferen-

ças e semelhanças em relação às características acústicas típicas do gênero cisgênero correspondente, com variações específicas observadas entre homens trans e homens cis, bem como entre mulheres trans e mulheres cis.

4 Resultados e Discussão

Para uma melhor organização, os resultados são apresentados de acordo com os objetivos estabelecidos, começando com os dados obtidos por meio do julgamento perceptivo-auditivo feitos pelos juízes. Em seguida, são apresentados os resultados da análise acústica dos sujeitos trans e cis no que diz respeito aos aspectos de natureza vocal e prosódica investigados. Por fim, são expostas as comparações entre os sujeitos transgênero e cisgênero.

4.1 Avaliação perceptivo-auditiva

4.1.1 Julgamento do gênero do falante

Os dados obtidos no julgamento perceptivo-auditivo do gênero do falante estão apresentados na Tabela 1.

Tabela 1 – Atribuição de gênero feita pelos juízes aos sujeitos trans no julgamento perceptivo-auditivo

Sujeitos	Voz masculina %	Voz feminina %	Indefinido %	Considerado	Final %
HT1	33,3	0	66,67	Indefinido	Voz masculina = 60%
HT2	100	0	0	Voz masculina	
HT3	33,33	66,67	0	Voz feminina	
HT4	100	0	0	Voz masculina	
HT5	100	0	0	Voz masculina	
MT1	0	100	0	Voz feminina	Voz feminina = 80%
MT2	0	100	0	Voz feminina	
MT3	0	66,67	33,33	Voz feminina	
MT4	66,67	0	33,33	Voz masculina	
MT5	33,33	66,67	0	Voz feminina	

Fonte: Elaboração própria.

Legenda: HT1=homem trans 1, HT2= homem trans 2, HT3= homem trans 3, HT4= homem trans 4, HT5= homem trans 5, MT1= mulher trans 1, MT2= mulher trans 2, MT3= mulher trans 3, MT4= mulher trans 4, MT5= mulher trans 5, Indefinido= voz masculina/voz feminina

Somam-se aos dados da Tabela 1, 100% de concordância entre os juízes leigos no julgamento da voz do HC como voz masculina e da MC como voz feminina. A partir desse dado, tem-se que os juízes perceberam, a partir do julgamento da voz, o gênero dos sujeitos cisgênero (HC e MC) como correlatos ao sexo de nascimento.

Em contrapartida, os mesmos juízes julgaram, preponderantemente, as vozes de sujeitos trans de acordo com o gênero declarado, ou seja, conforme mostra os dados da Tabela 1, as vozes de HT foram julgadas com 60% de concordância como vozes masculinas e as vozes de MT com 80% de concordância, como vozes femininas.

Este dado é discordante de estudo com a fala de MT brasileiras que tiveram suas vozes identificadas como femininas em pouco mais da metade das que foram julgadas por juízes leigos (Schmidt, 2018). A percepção da voz como sendo feminina foi discutida em diversos estudos e os resultados são controversos, pois alguns apontaram que tal percepção estaria atrelada à medida de f_0 (Gelfer; Bennett, 2013; Gelfer; Schofield, 2000) e, ainda, que quanto maior o valor de f_0 , maior a chance de uma voz ser julgada como feminina (Munson, 2007; Owen; Hancock, 2010), enquanto outros resultados de pesquisa apontaram que não só a f_0 determina o julgamento da voz como feminina, mas que outros aspectos da emissão da fala deveriam ser considerados em novos estudos (Hillenbrand; Clark, 2009). Dahl e Mahler (2020) encontraram, por exemplo, a relação da f_0 e a intensidade vocal, na percepção da feminilidade do falante.

No que diz respeito à fala de HT, os resultados de pesquisa considerando a fala de homens transexuais abordam, principalmente, processos terapêuticos, cirúrgicos ou fonoaudiológicos adotados por esse grupo e destacam como esses recursos foram avaliados por meio da aplicação de instrumentos de autoavaliação e autossatisfação. A esse respeito, McNeill *et al.* (2006) demonstraram que a satisfação com a própria voz por HT após terapia fonoaudiológica não se relacionou apenas com a medida de f_0 .

É válido destacar que, neste estudo, não foi feito um controle de quais sujeitos passaram por tratamentos cirúrgico, hormonal ou fonoaudiológico para mudança das vozes. De todo modo, para a avaliação perceptivo-auditiva, essa limitação do estudo não afeta os resultados, uma vez que o principal objetivo dessa análise era saber qual(is) gênero(s) seria(m) atribuído(s) aos sujeitos trans influentes na mídia digital a partir do julgamento perceptivo-auditivo feito pelos juízes no que diz respeito à voz desses sujeitos, independentemente se esses sujeitos passaram ou não por procedimentos clínicos e/ou cirúrgicos.

Diante do exposto, conforme apontado por outros estudos já realizados, entende-se aqui que outros recursos vocais e prosódicos da fala devem ser considerados para melhor compreensão sobre como são julgadas as vozes de MT e HT. Entoação, velocidade de fala, intensidade vocal (Houle; Levi, 2021) e até mesmo vocabulário (Neumann; Welzel, 2004) são citados como recursos a serem considerados nesse julgamento. Neste artigo, centramo-nos em analisar também a marcação de fronteiras e proeminências prosódicas.

4.1.2 Julgamento de fronteiras prosódicas e marcação de proeminências

Os dados referentes ao julgamento perceptivo-auditivo de fronteiras prosódicas e proeminências estão organizados na Tabela 2 e nos gráficos 1 e 2. Fronteiras e proeminências identificadas pelos juízes foram contabilizados e, posteriormente, divididos pelo número de palavras da amostra de cada sujeito. Esse procedimento foi adotado de modo a obter um índice da frequência desses recursos prosódicos em relação à duração total das amostras de fala analisadas para, assim, viabilizar a comparação dos sujeitos entre si no que diz respeito à expressividade marcada pelas proeminências e à segmentação do discurso conforme o julgamento dos juízes. Optou-se por esse procedimento já que as amostras de fala não foram controladas e, portanto, não foram produzidas a partir de um texto comum a partir do qual as proeminências e as fronteiras prosódicas poderiam ter sido identificadas a partir da estrutura informacional⁵ para, posteriormente, caracterizar a variação entre os sujeitos no uso desses recursos. Após a obtenção desse índice, para a análise das fronteiras e das proeminências identificadas na avaliação perceptivo-auditiva em cada grupo de sujeitos trans, calculou-se a média e o desvio padrão.

A Tabela 2 apresenta a média e o desvio padrão (DP) do índice de fronteiras e de proeminências extraídas de cada grupo de sujeitos trans e o índice desses eventos para HC e MC.

Tabela 2 – Relação de fronteiras e proeminências identificadas por número de palavras produzidas nas amostras de fala de MT, MC, HT e HC

Parâmetros	HT		HC	MT		MC
	Média do Índice	DP		Média do Índice	DP	
Fronteiras	0,22	0,12	0,15	0,22	0,10	0,21
Proeminência	0,10	0,03	0,18	0,19	0,02	0,34

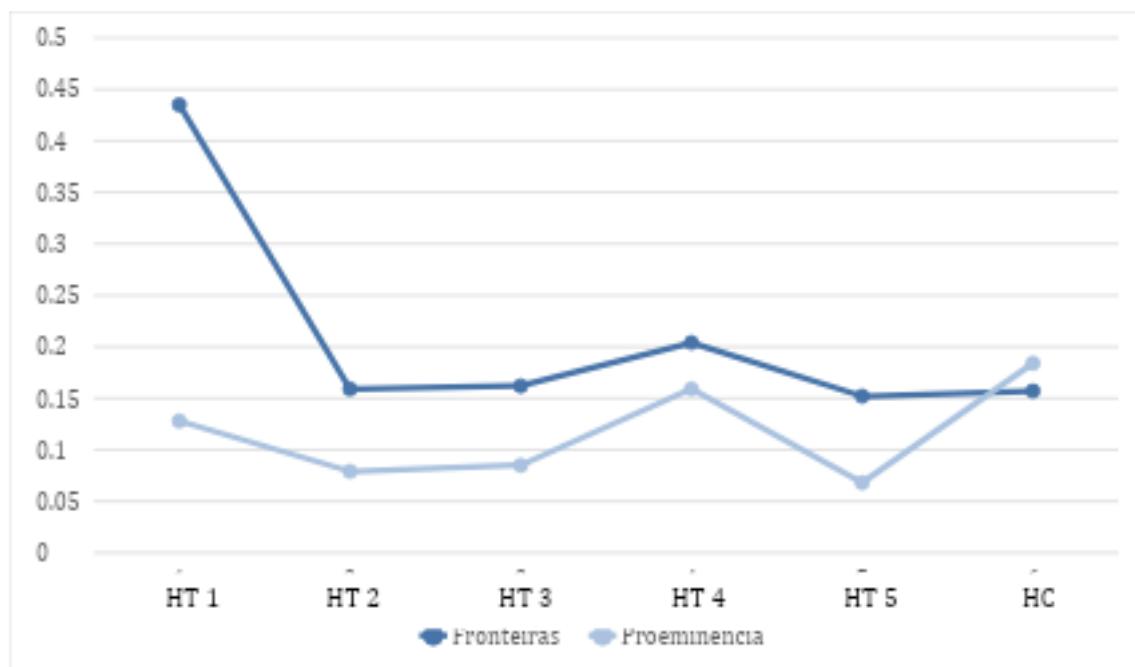
Fonte: Elaboração própria.

Legenda: HT = homem trans, HC= homem cis, MT=mulher trans, MC= mulher cis.

Por sua vez, os Gráficos 1 e 2 apresentam a variação dos índices de fronteiras e de proeminências de cada sujeito trans e do sujeito cis.

⁵ Da literatura, sabe-se que as proeminências dependem da estrutura informacional. No entanto, considerando que as amostras de fala analisadas não foram controladas por terem sido extraídas de mídia social, optou-se pelo cálculo descrito como maneira de normalizar as variações inerentes às amostras de fala analisadas.

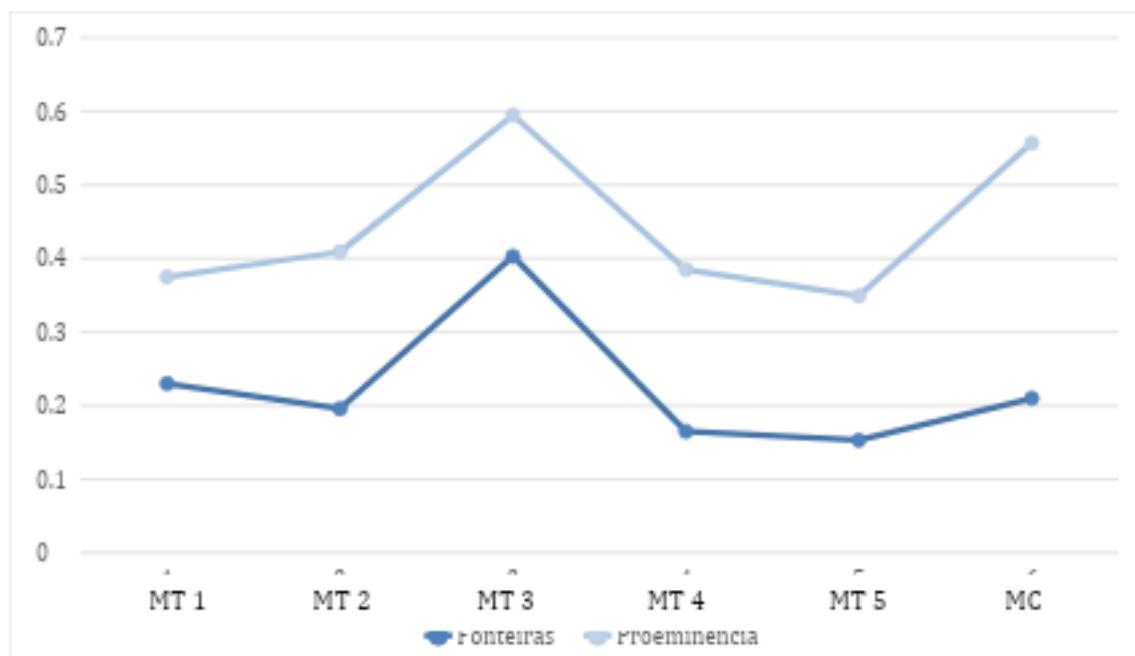
Gráfico 1– Variação dos índices de fronteiras e de proeminências identificadas por sujeito no grupo HT e no HC



Fonte: Elaboração própria.

Legenda: HT1=homem trans 1, HT2= homem trans 2, HT3= homem trans 3, HT4 = homem trans 4, HT5 = homem trans 5, HC=homem cis.

Gráfico 2 – Variação dos índices de fronteiras e de proeminências identificadas por sujeito no grupo MT e no MC



Fonte: Elaboração própria.

Legenda: MT1=mulher trans 1, MT2= mulher trans 2, MT3 = mulher trans 3, MT4 = mulher trans 4, MT5 = mulher trans 5, MC = mulher cis.

Observa-se na Tabela 2 que HT apresentou maior índice de fronteiras prosódicas identificadas que HC, enquanto o grupo de MT apresentou índice similar de fronteiras comparado a MC. No que tange aos índices de identificação de proeminências, os grupos HT e MT apresentaram índices menores que HC e MC respectivamente; entretanto MT comparada a MC apresentou diferença mais acentuada (o índice de MC em relação a MT é o dobro do índice de comparação entre HC e HT). No entanto, quando aprofundamos a interpretação do julgamento realizado pelos juízes em relação a fronteiras e proeminências, considerando as diferenças entre sujeitos trans no interior de cada grupo comparado aos sujeitos cis, encontramos variações notáveis entre eles.

Em relação às fronteiras, observa-se que HT1 apresentou maior índice que os outros sujeitos do grupo HT, nos quais os índices foram próximos entre si e, ainda, similares ao HC (Gráfico 1). De forma parecida, a MT3 exibiu uma frequência mais elevada do índice de fronteiras prosódicas identificadas quando comparada à MC e às outras MT, ou seja, com exceção de MT3, os demais sujeitos do grupo MT e também MC apresentaram similaridade no índice de fronteiras prosódicas identificadas (Gráfico 2).

Em relação à proeminência, na fala dos sujeitos HT, palavras proeminentes foram identificadas com menor índice em comparação à fala de HC, ainda que o índice de HT4 tenha sido o mais próximo em relação a HC (Gráfico 1). Por sua vez, o julgamento da proeminência na fala de MT foi similar em quatro das cinco mulheres trans que formam o grupo, evidenciando uma consistência marcante entre elas, embora o índice de identificação na fala delas tenha sido menor em relação a MC. Entretanto, observa-se que a fala da participante MT3 obteve um índice de identificação de fronteiras mais elevado quando comparada às demais e foi aquela que mais se aproximou de MC (Gráfico 2).

De maneira geral, a análise dos índices de identificação de fronteiras e proeminência mostra que, apesar de esses índices apontarem para a possibilidade de identificação de estilos de fala de sujeitos particulares, como é o caso de HT1 e MT3, eles sugerem, com base no julgamento realizado na avaliação perceptivo-auditiva, que a fala dos sujeitos trans pode apresentar características similares enquanto um grupo, diferenciando-se ou não do sujeito cis, pois (i) no caso das fronteiras, a fala de HT foi, em geral, julgada como próxima a HC e, igualmente, a fala de MT foi, em geral, julgada como próxima a MC, notando-se, assim, uma aproximação do estilo de fala de sujeitos trans com o estilo de fala do gênero declarado por esses sujeitos no que tange à delimitação de fronteiras prosódicas; e (ii) no caso das proeminências, tanto a fala de HT quanto a fala de MT foi, em geral, julgada como diferente em relação a HC e MC, respectivamente, pois tanto no grupo HT quanto no grupo MT, em geral, a proporção de palavras proeminentes identificadas foi menor em relação aos sujeitos cis, notando-se, assim, um distanciamento do estilo de fala dos grupos de sujeitos trans com o estilo de fala do gênero declarado no que tange a marcação de proeminências.

Interpretamos que os resultados ora apresentados parecem sugerir um rompimento com uma imagem estereotipada da fala dos sujeitos trans. Imagem essa socialmente construída com base na normatividade binária e de acordo com a qual a fala da população trans torna-se, geralmente, associada ao exagero ou à artificialidade num processo discursivo que tem como consequência caracterizá-la como anormal ou desviante. Os dados obtidos na avaliação perceptivo-auditiva sugerem romper com essa imagem na medida em que mostram, por um lado, tendência de aproximação do estilo de fala dos sujeitos trans com o estilo de fala dos sujeitos cis na marcação de fronteiras e, ainda, por outro lado, não mostram exagero no

uso de proeminências dos sujeitos trans em relação aos sujeitos cis, uma vez que, comparativamente, a fala trans foi avaliada com menos pontos de marcação de proeminência do que a fala cis. Assim, considerando que a marcação de proeminência pode ser considerada um importante recurso expressivo nas situações de interação verbal, o distanciamento do estilo de fala dos sujeitos trans em relação aos sujeitos cis nas proeminências identificadas parece por em cheque a imagem estereotipada da fala dos sujeitos trans, pois a fala desses sujeitos foi considerada, em geral, menos proeminente marcada na amostra analisada do que a fala dos sujeitos cis.

4.2 Análise acústica

4.2.1 Análise da Frequência fundamental

Os valores obtidos referentes à frequência fundamental da fala para homens e mulheres trans são apresentados na Tabela 3.

Tabela 3 – Valores de f_0 da fala de HT e MT

Parâmetro f_0 (Hz)	Falantes	1	2	3	4	5	Média	DP
	HT	138,427	152,799	133,202	140,652	117,145	136,445	12,962
MT	135,586	160,658	215,485	134,097	220,011	173,167	42,072	

Fonte: Elaboração própria.

Legenda: HT= Homens Trans, MT= Mulheres Trans, DP= Desvio Padrão

Deve-se acrescentar à leitura da tabela que HC apresentou o valor de f_0 igual a 162,062 Hz e MC apresentou valor igual a 239,137 Hz.

Ao comparar o valor médio de f_0 do grupo de HT com o valor de HC, observa-se que os valores de HT estão abaixo do valor de HC, com uma diferença de 25,617 Hz, demonstrando o uso de um *pitch* mais grave pelo grupo de HT. Uma revisão de literatura apresentada em Spazzapan *et al* (2019), abrangendo estudos de várias nacionalidades e faixas etárias de 18 a 50 anos, constatou que a f_0 variou entre 190 e 251 Hertz para mulheres e entre 112 e 141 Hz para homens. Assim, nos dados encontrados neste estudo, conforme os procedimentos metodológicos adotados, o HC apresentou f_0 fora do esperado para homens, enquanto o grupo HT apresentou valores dentro do intervalo reportado pela literatura. Vale lembrar que o valor de f_0 do HC está na faixa reconhecida como neutra, que pode ser entendida como favorável para a identificação da voz tanto como masculina, quanto como feminina (Hardy *et al.*, 2018).

Em relação ao grupo MT, é possível verificar que há uma maior discrepância com o valor de f_0 da MC, com diferença de 65,97 Hz, demonstrando que o grupo MT apresentou o *pitch* mais grave em relação a MC, sendo a média do grupo dentro da faixa de frequência esperada para vozes femininas, entre 190 Hz e 251 Hz (Spazzapan *et al.*, 2019). Porém, há valores de sujeitos que, individualmente, não estão dentro dessa média. Tais resultados apontam

que os sujeitos trans podem adaptar f_0 de forma a se aproximar dos valores típicos do gênero declarado, ainda que nem sempre aconteça.

Observando os resultados da medida de f_0 dos participantes transgênero e os da avaliação perceptivo-auditiva do gênero do falante (apresentada em 4.1.1), pode-se levantar o questionamento sobre os motivos que teriam motivado as vozes do grupo HT terem sido julgadas pelos juízes como masculinas na proporção de 60% se o valor de f_0 encontrado na análise acústica para todos os HT esteve dentro da faixa de frequência esperada para vozes masculinas. Uma hipótese que se levanta aqui é a de que quem avaliou as falas teve apoio no estilo de fala marcado pelos aspectos prosódicos.

Um fato intrigante que deve ser estudado de forma mais aprofundada é que HT1 e HT3 apresentaram vozes com f_0 por volta de 130 Hz, uma faixa de frequência de voz masculina, entretanto o julgamento perceptivo-auditivo de HT1 foi de voz indefinida e de HT3, feminina. Nestes julgamentos, a definição do gênero não foi apoiada na f_0 . Num estudo de revisão de literatura e meta análise foi observado que 58,4% de estudos sobre atribuição de gênero por avaliação perceptivo-auditiva, tiveram seus resultados explicados por outros fatores da comunicação (Hardy *et al.*, 2018). Wolfe *et al* (1990), por sua vez, concluíram que uma medida limite de f_0 na voz do transgênero para ser julgada como feminina seria 155 Hz, mas afirmam ainda que a variação de frequência fundamental, que caracterizaria padrões de entoação com picos e vales menos acentuados, poderia ser um importante padrão na percepção na voz do HT.

Entretanto, as vozes do grupo MT, ainda que tenham apresentado valor de f_0 mais baixo em relação a MC, foram privilegiadamente julgadas como vozes femininas (80%), fato que pode ser explicado pela faixa de frequência considerada neutra (Hardy *et al.*, 2018). Desse modo, os resultados encontrados corroboram estudos que apontaram que valores de f_0 a partir de 155 Hz (Wolfe *et al.*, 1990) ou de 165 Hz podem ser julgados e associados a vozes femininas na fala de MT, ou ainda, valores maiores que 180 Hz (Gorham-Rowan, Morris, 2006). No grupo MT, a única voz que foi julgada como masculina (MT4) apresentou medida acústica de 134 Hz, valor que, por estar abaixo do valor de corte apresentado pela literatura para reconhecimento de uma voz feminina, parece justificar tal julgamento. Entretanto, é relevante destacar que uma das vozes da amostra (MT1) apresentou medida similar com valor de 135,5 Hz e, ainda assim, foi julgada como feminina. Em conjunto, os resultados obtidos no cotejamento da análise acústica com o julgamento perceptivo-auditivo propõem a reflexão de que a percepção do outro sobre o gênero do falante, ao invés de recuperar unicamente um padrão acústico de reconhecimento de voz, parece considerar informações de diferentes naturezas, tais como informações acerca do estilo de fala marcado por aspectos prosódicos conforme passamos a apresentar a seguir.

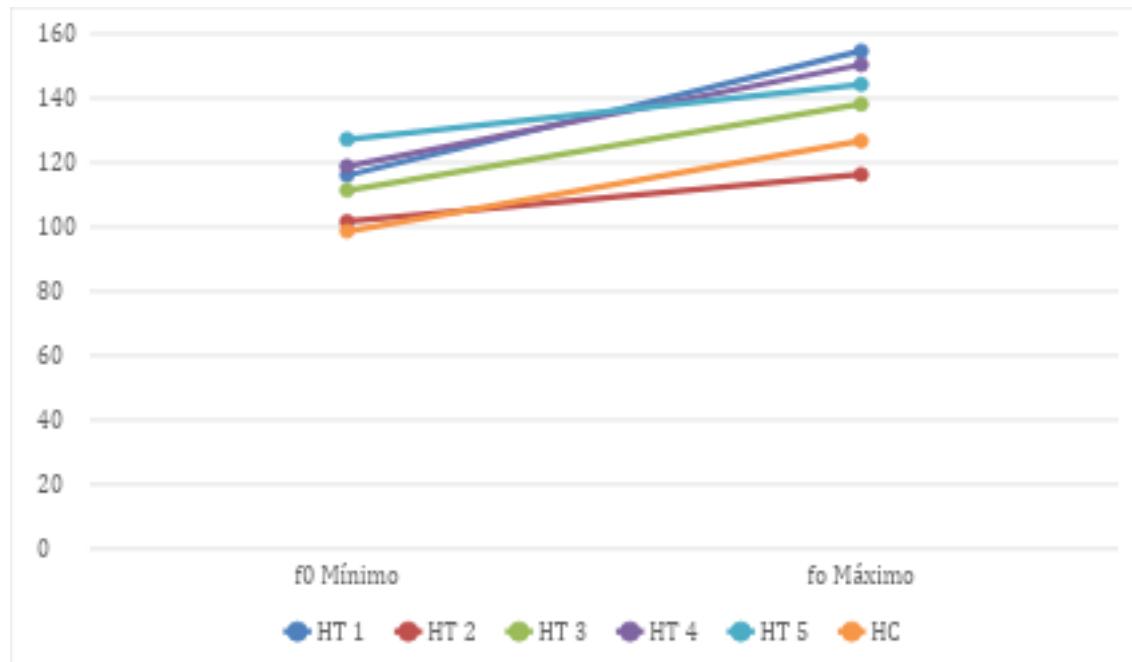
4.2.2 Caracterização acústica das fronteiras prosódicas

Conforme descrito na metodologia, a análise acústica das fronteiras prosódicas que foram identificadas perceptivo-auditivamente pelos juízes foi realizada pela mensuração de pausas e dos valores de f_0 mínima e máxima da sílaba tônica da palavra produzida imediatamente anterior à fronteira. A seguir, apresentamos os resultados das medidas mensuradas.

4.2.2.1 Análise da variação de frequência fundamental

O Gráfico 3 apresenta os valores médios de f_0 mínima e máxima mensurados na fala de HT e HC. Por sua vez, o Gráfico 4 apresenta os valores mensurados na fala de MT e MC.

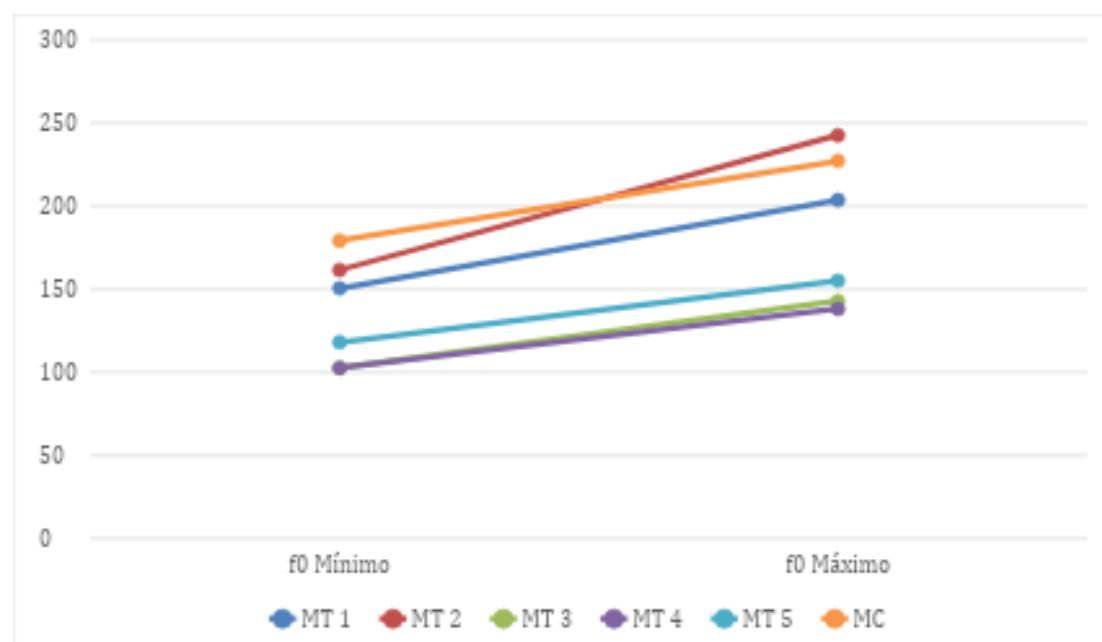
Gráfico 3 – Média de f_0 mínima e máxima da palavra que antecede as fronteiras identificadas na fala de HT e HC



Fonte: Elaboração própria.

Legenda: HT1= Homens Trans 1, HT2= Homem Trans 2, HT3= Homem Trans 3, HT4= Homem Trans 4; HT5= Homem Trans 5; HC= Homem Cis

Gráfico 4 – Média das medidas de f_0 mínima e máxima da palavra que antecede as fronteiras identificadas na fala de MT e MC



Fonte: Elaboração própria.

Legenda: MT1= Mulher Trans, MT2= Mulher Trans 2, MT3= Mulher Trans 3, MT4= Mulher Trans 4, MT5= Mulher Trans 5, MC= Mulher Cis

Os gráficos 3 e 4 indicam que as fronteiras identificadas pelos juízes foram marcadas na fala de todos os sujeitos, independentemente do grupo, por variação de f_0 , a qual foi ilustrada nos gráficos pela diferença entre os seus valores mínimo e máximo na palavra que antecedeu a fronteira. Assim, os dados encontrados estão em consonância com resultados já apresentados para o Português Brasileiro, obtidos por meio de análise inferencial, diferentemente dos nossos, de que fronteiras prosódicas como as fronteiras de frase entoacional são marcadas por variações de f_0 , as quais são relevantes para a identificação das fronteiras pelos ouvintes (Serra, 2009; Soncin; Tenani, 2016).

O Gráfico 3, no entanto, permite observar ainda, no que diz respeito à comparação entre HT e HC, que a faixa de f_0 usada pelo grupo HT foi, em geral, maior do que a usada pelo HC. Esse dado é relevante, pois sugere que as fronteiras foram, em geral, marcadas acusticamente pelos HT com uma faixa de frequência mais alta, o que indicaria eventos tonais de fronteira marcados de forma mais aguda em relação a HC. Em contrapartida, o gráfico 4 mostra que o grupo MT, em geral, empregou uma faixa de f_0 mais baixa na marcação das fronteiras em comparação com MC, o que indicaria, na direção oposta ao que se observou para HT, eventos tonais de fronteira marcados de forma mais grave em relação a MC.

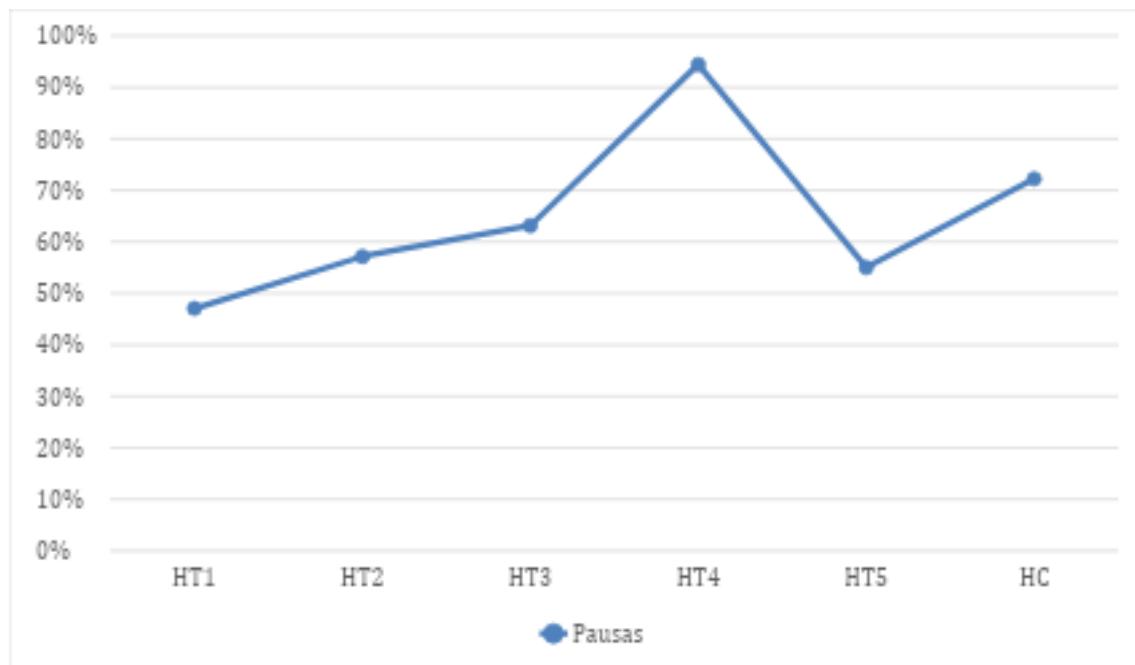
Tais resultados mostram, assim, que na marcação de fronteiras os sujeitos trans se diferenciam dos sujeitos cis. Ou seja, na sinalização das rupturas prosódicas, observam-se indícios que diferenciam HT em relação a HC por meio de fronteiras marcadas de forma mais aguda e indícios que diferenciam MT em relação a MC com fronteiras marcadas de forma mais grave. Diferentemente do que se observa em outras medidas, na faixa de f_0 com a qual as fronteiras são marcadas, sujeitos trans parecem se mostrar como um grupo outro em relação àquele do gênero declarado, o que indicaria uma identidade de grupo singular, frente ao deslocamento em relação aos sujeitos cis. Destaca-se aqui o ineditismo desse achado.

A esse respeito, retomamos a proposição de Ribeiro e Sobral (2021), segundo a qual novas representações identitárias, ou seja, novas ordens se impõem - e essas se marcam no discurso -, quando o deslocamento do sujeito em relação a seu grupo de pertença encontra o vazio como espaço simbólico. Sob essa perspectiva, o distanciamento dos sujeitos trans em relação ao sujeitos cis que se mostra na marcação de fronteiras prosódicas, marca os limites da relação entre sujeito trans e sujeito cis na identidade de gênero, pois, embora os sujeitos trans se identifiquem como homem ou como mulher de acordo com o gênero que autodeclararam, esses sujeitos também se identificam numa relação dialética com os homens e mulheres cis, pois esse sujeitos se identificam como trans, ou seja, não se trata de “sou homem” ou “sou mulher”, trata-se de se identificaram nas mídias sociais, como “sou homem trans” e “sou mulher trans”: uma nova ordem, portanto, nova representação identitária, fruto do vazio identificado com espaço simbólico na relação que estabelecem com homens e mulheres cis.

4.2.2.2 Análise da produção de pausas

No que diz respeito à análise da pausa como marcador de fronteira, os Gráficos 5 e 6 apresentam o percentual das fronteiras identificadas que foram marcadas acusticamente com pausa pelos sujeitos dos diferentes grupos que compõem a amostra.

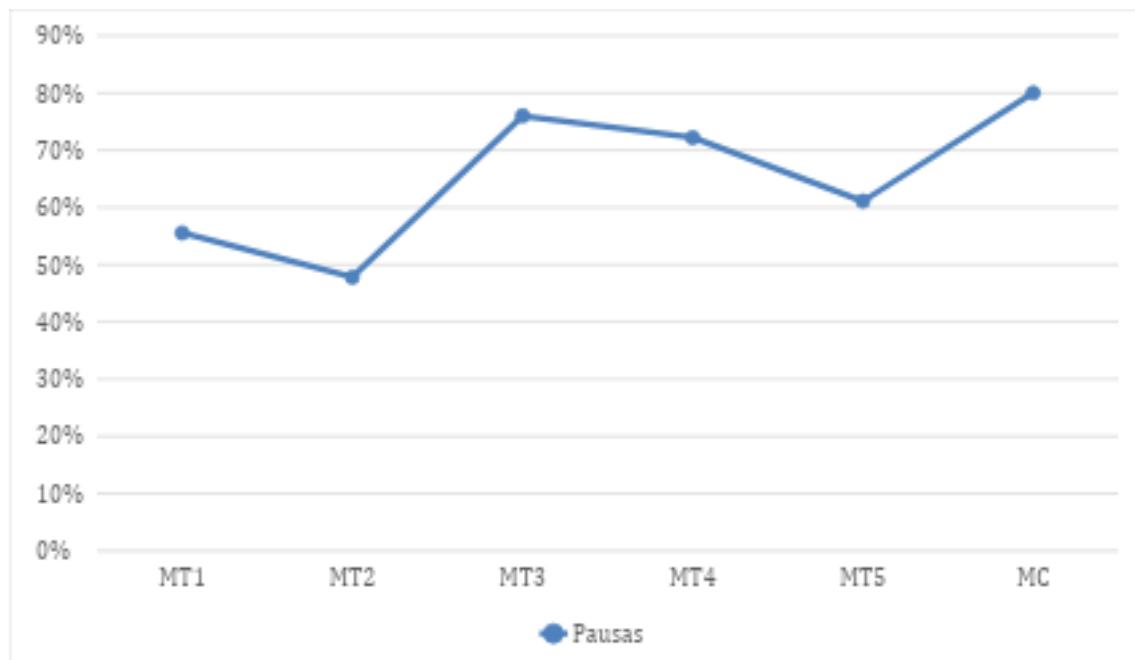
Gráfico 5 – Percentual de fronteiras prosódicas marcadas por pausas em HT e HC



Fonte: Elaboração própria.

Legenda: HT1= Homens Trans 1, HT2= Homem Trans 2, HT3= Homem Trans 3, HT4= Homem Trans 4; HT5= Homem Trans 5; HC= Homem Cis

Gráfico 6 – Percentual de fronteiras prosódicas marcadas por pausas em MT e MC



Fonte: Elaboração própria.

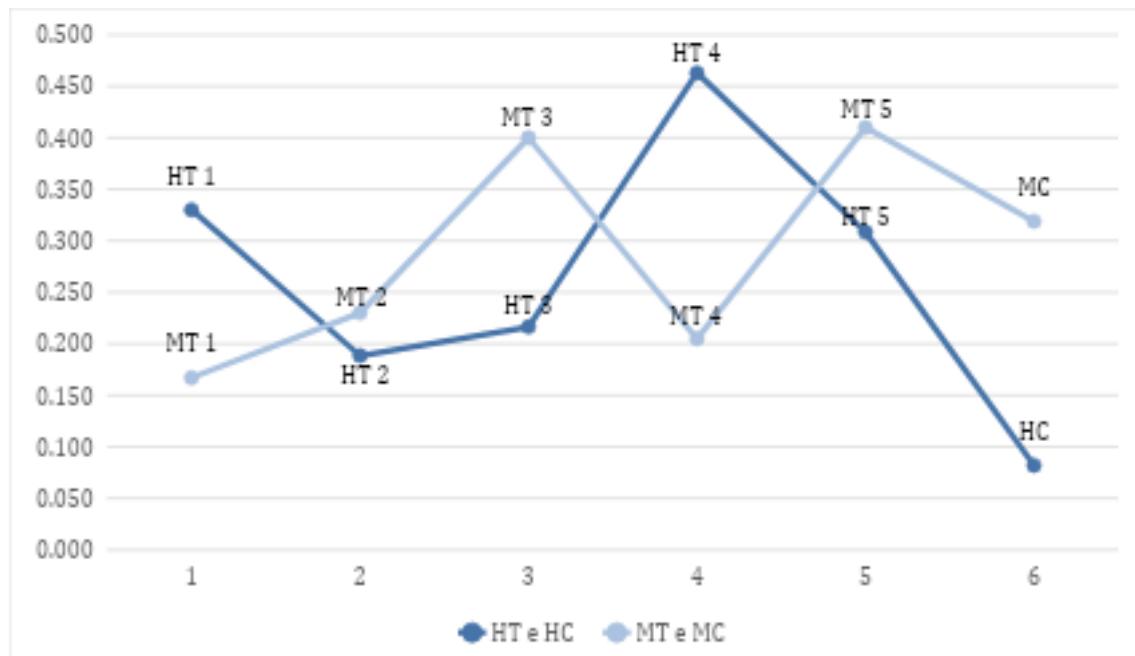
Legenda: MT1= Mulher Trans, MT2= Mulher Trans 2, MT3= Mulher Trans 3, MT4= Mulher Trans 4, MT5= Mulher Trans 5, MC= Mulher Cis

Os gráficos 5 e 6 mostram que, para todos os sujeitos avaliados, independentemente de sua afiliação a um grupo específico, as fronteiras identificadas pelos juízes na avaliação perceptivo-auditiva não foram necessariamente marcadas por pausas no sinal acústico. No caso de HT, o percentual de presença de pausas nas fronteiras identificadas oscilou entre 50% e 60% (Gráfico 5) para a maioria dos sujeitos (com média igual a 63,36%), tendo sido observado o maior percentual para HT4 (mais que 90%) e um percentual de aproximadamente 70% para HC. No caso de MT, o percentual de presença de pausa foi mais variável entre os sujeitos (com média igual a 62,545) e, ainda, oscilou em intervalo de maior diferença (entre 50 e 80%), sendo o percentual de MC aquele de maior valor (Gráfico 6).

Esses resultados, por um lado, estão alinhados com achados de estudo anterior realizado com estatística inferencial que mostraram não haver necessária identidade entre a percepção de fronteira prosódica e a presença de pausa no sinal acústico (Soncin; Tenani; Berti, 2017; 2019), uma vez que variações de f_0 no sinal acústico, mesmo quando não combinadas a pausas, levam a identificação de fronteiras prosódicas e podem ser julgadas como pausas perceptuais. Por outro lado, os resultados apontam para variações entre sujeitos em ambos os grupos, o que nos faz interpretar a marcação de fronteiras prosódicas por pausas como um fenômeno variável mais relacionado à dinâmica da fala e menos suscetível ao estilo de fala de grupos de sujeitos.

A variação entre sujeitos no interior dos grupos é observada também quando se considera a duração da pausa produzida. O Gráfico 7 apresenta a média de duração das pausas que foram utilizadas como marcadores de fronteira na fala dos diferentes sujeitos.

Gráfico 7 – Média de duração em segundos das pausas produzidas em fronteiras prosódicas por sujeito de cada grupo e pelos sujeitos cis



Fonte: Elaboração própria.

Legenda: HT1=homem trans 1, HT2=homem trans 2, HT3=homem trans 3, HT4 = homem trans 4, HT5 = homem trans 5, HC=homem cis, MT1=mulher trans 1, MT2= mulher trans 2, MT3 = mulher trans 3, MT4 = mulher trans 4, MT5 = mulher trans 5, MC = mulher cis

Do Gráfico 7 destaca-se que não houve regularidade de duração de pausa entre os sujeitos analisados e, esclarece-se que, considerando a ausência de estudos anteriores que tenham considerado essa variável na fala dos sujeitos trans, é difícil esboçar uma comparação entre sujeitos trans e cis. Entretanto, ressalta-se que a alta variação duracional observada pode sugerir a manipulação das pausas pela edição dos vídeos para adaptação às plataformas digitais. Isso implica que as pausas podem ter sido manipuladas ou ajustadas durante o processo de edição, o que pode ter impactado na interpretação das fronteiras prosódicas nas gravações analisadas. Essa observação ressalta a importância de considerar possíveis intervenções na análise de dados audiovisuais, especialmente em contextos digitais.

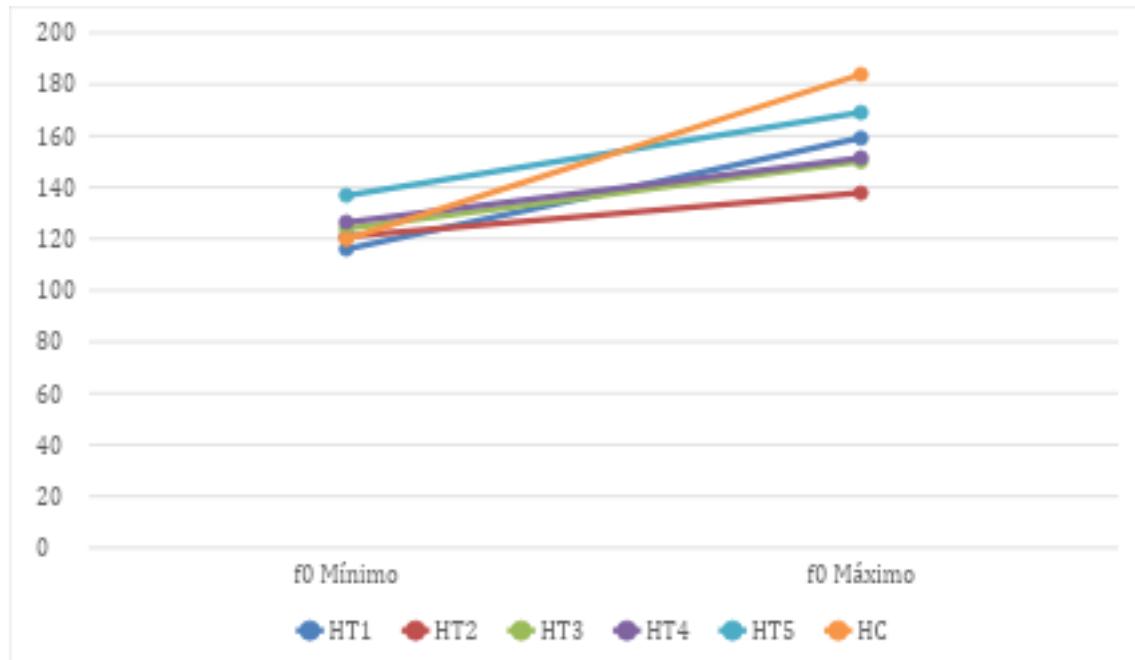
4.2.3 Caracterização acústica das Proeminências

Finalmente, apresentam-se os dados obtidos na análise acústica das proeminências. Acusticamente, essas proeminências foram analisadas pela mensuração de f_0 mínima e máxima da sílaba tônica da palavra identificada como proeminente, bem como pela mensuração da intensidade mínima e máxima. Apresenta-se, inicialmente, a análise dos valores de f_0 .

4.2.3.1 Análise da variação de frequência fundamental

As médias das medidas mínimas e máximas de f_0 na sílaba tônica da palavra proeminente identificada pelos juízes estão apresentadas nos Gráficos 8 e 9.

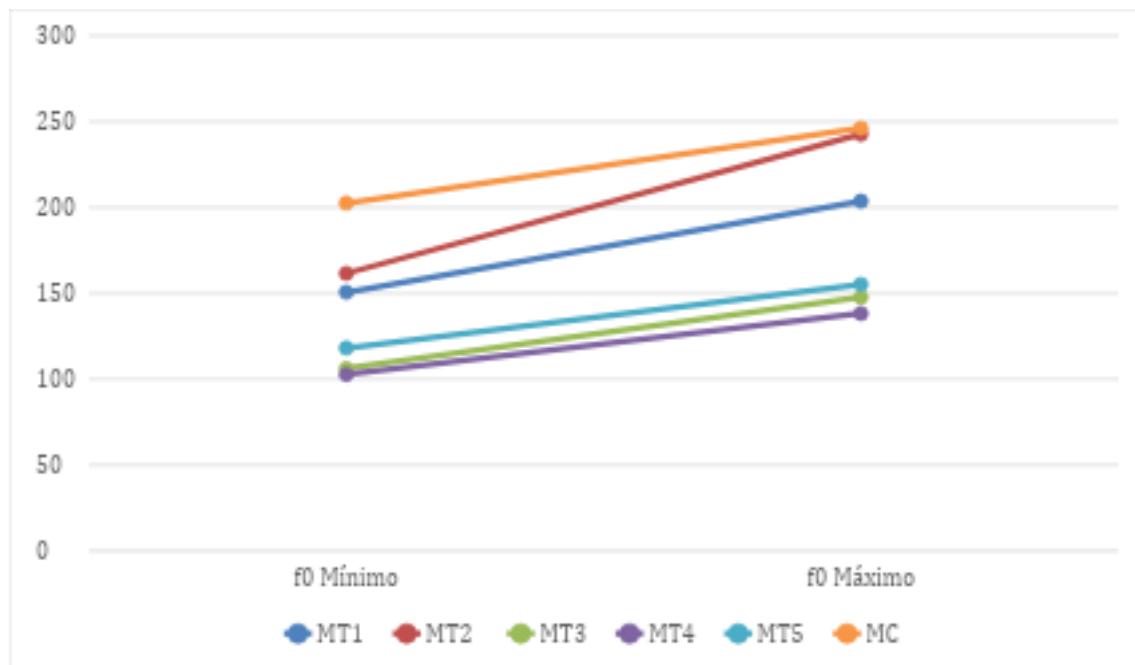
Gráfico 8 – Médias de f_0 mínima e máxima na marcação de proeminência por HT e HC



Fonte: Elaboração própria.

Legenda: HT1=homem trans 1, HT2= homem trans 2, HT3= homem trans 3, HT4 = homem trans 4, HT5 = homem trans 5, HC=homem cis

Gráfico 9 – Médias de f_0 mínima e máxima na marcação de proeminência por MT e MC



Fonte: Elaboração própria.

Legenda: MT1=mulher trans 1, MT2= mulher trans 2, MT3 = mulher trans 3, MT4 = mulher trans 4, MT5 = mulher trans 5, MC = mulher cis

A leitura dos gráficos 8 e 9 permite observar, em primeira instância, que as proeminências identificadas na avaliação perceptivo-auditiva foram marcadas acusticamente por diferença nos valores de f_0 mínima e máxima de, ao menos, 30 Hz. Assim, os dados indicam pontos de congruência entre a identificação da proeminência e a variação nos valores de f_0 , corroborando os dados da literatura linguística que apontam f_0 como o correlato acústico mais robusto para a marcação de proeminência (Barbosa; Madureira, 2015, Gussenhoven, 2006; Moraes, 2009; Terken; Hermes, 2000, entre outros).

Em segunda instância, a leitura dos gráficos permite ainda observar comportamentos diferentes marcados acusticamente nas proeminências quando se comparam sujeitos trans e cis nos diferentes grupos contemplados.

No caso do grupo HT (Gráfico 8), a diferença entre mínima e máxima de f_0 na marcação de proeminência foi menor em relação ao sujeito cis, haja vista o valor mais alto de f_0 máxima identificado na amostra de fala desse sujeito. Como efeito, os valores sugerem que HC explorou de modo mais marcado as proeminências em seu estilo de fala quando comparado a todos os sujeitos HT. Ou seja, os valores na análise acústica parecem indicar tendência do grupo HT a uma fala menos marcada em relação a HC, no sentido de oscilações menos amplas de f_0 para a marcação de proeminências, o que daria subsídios para se interpretar a fala de HT como distanciada de uma imagem ligada ao exagero, geralmente associada à fala da população trans.

No caso do grupo MT (Gráfico 9), observa-se também comportamentos que distinguem os sujeitos trans desse grupo em relação a MC. No entanto, os valores que permitem essa observação não dizem respeito à diferença entre mínimo e máximo de f_0 , mas sim às

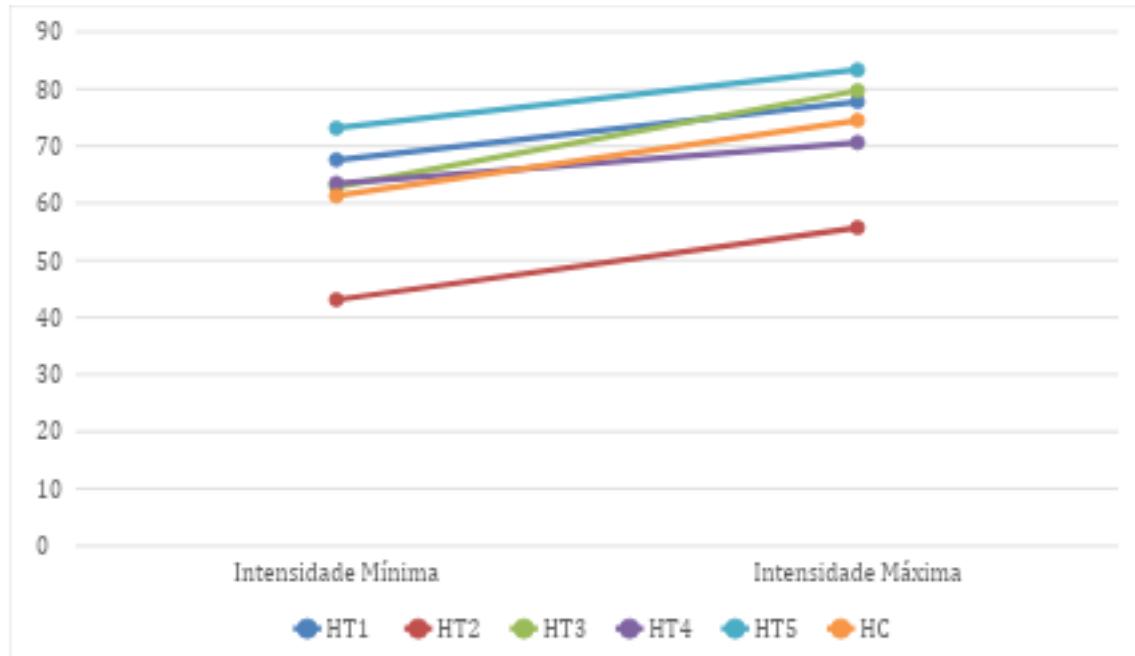
diferentes faixas de f_0 utilizadas para marcar a proeminência. Enquanto MC utilizou uma faixa entre 200 e 250 Hz, os sujeitos do grupo MT tenderam a utilizar uma faixa mais baixa, entre 100 e 200 Hz. Como efeito, as proeminências do grupo MT indicam uma distinção acústica para marcar a proeminência por meio de uma faixa de frequência mais grave em relação à MC, o que pode ser interpretado novamente como uma tendência a um estilo de fala menos caricato e, portanto, com proeminências marcadas por meio de eventos tonais mais graves.

Esses resultados da análise acústica das proeminências são congruentes com os resultados referentes ao julgamento perceptivo-auditivo das proeminências, na qual também se observou que o recurso de proeminência foi menos marcado pelos grupos trans (cf. página 17 deste artigo) do que pelos sujeitos cis. Assim, em ambas as análises, têm-se indícios que permitem questionar a imagem não natural ou caricatural geralmente associada pelo senso comum à fala da população trans, já que, na amostra analisada, os resultados apresentaram índice menor de proeminências identificadas, menor variação e valores de parâmetros acústicos menos acentuados nessa população do que nos sujeitos cis. Na esteira de Ribeiro e Sobral (2021), podemos interpretar a diferenciação na fala dos sujeitos trans em relação aos sujeitos cis como um estilo de fala de um grupo que, discursivamente, tem inaugurado uma configuração outra, não exagerada ou caricata, como poderia se supor, distinta em relação aos sujeitos cis ainda que associada a ela, e comunicativamente eficiente nas mídias sociais.

4.2.3.2 Análise da Intensidade

No que diz respeito à intensidade, os gráficos 10 e 11 apresentam a média dos valores mínimo e máximo extraídos da sílaba tônica das palavras identificadas como proeminentes.

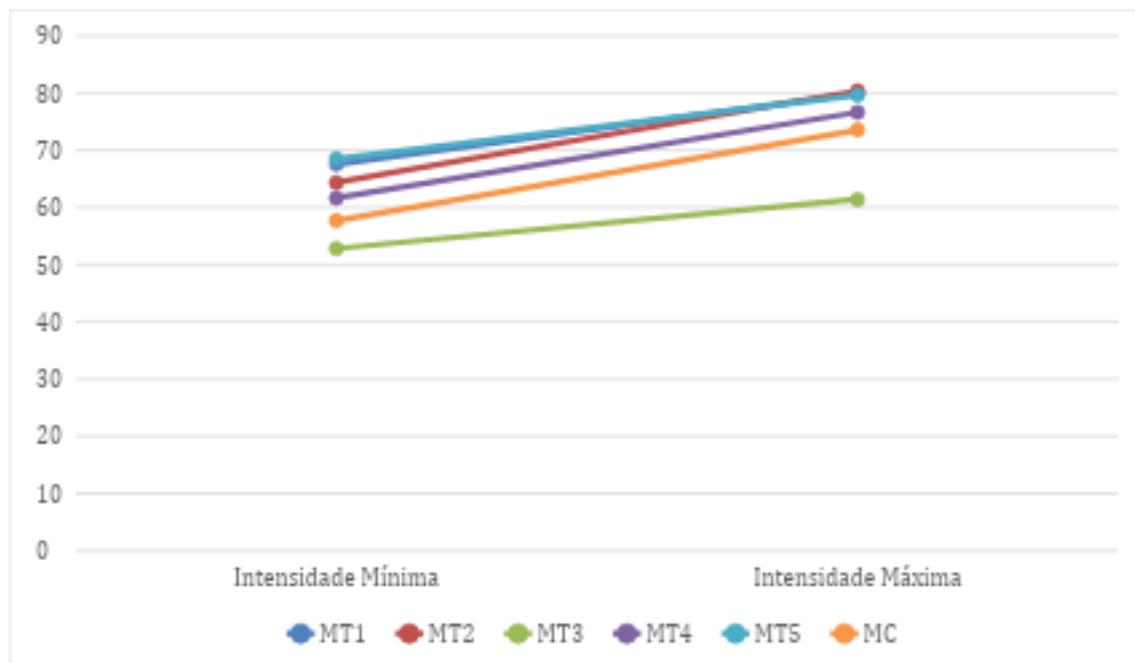
Gráfico 10 – Média de intensidade mínima e máxima na marcação de proeminências por HT e HC



Fonte: Elaboração própria.

Legenda: HT1=homem trans 1, HT2=homem trans 2, HT3=homem trans 3, HT4 = homem trans 4, HT5 = homem trans 5, HC=homem cis

Gráfico 11 – Média de intensidade mínima e máxima na marcação de proeminências por MT e MC



Fonte: Elaboração própria.

Legenda: MT1=mulher trans 1, MT2= mulher trans 2, MT3 = mulher trans 3, MT4 = mulher trans 4, MT5 = mulher trans 5, MC = mulher cis

Ambos os gráficos mostram que as proeminências identificadas foram marcadas por diferenças de intensidade mínima e máxima, ainda que essas diferenças não tenham sido de grande magnitude, já que não ultrapassaram 20dB. Essa pequena diferença, de certo modo, alinha-se descritivamente com estudos anteriores que não indicaram a intensidade como um parâmetro estatisticamente significativo para a marcação de proeminência e/ou não a consideram como um parâmetro robusto (Barbosa; Madureira, 2015; Gussenhoven, 2006; Moraes 2009; Terken; Hermes, 2000; entre outros).

Ademais, nos gráficos 10 e 11, pode-se observar uma variação pequena nos valores de intensidade mínima e máxima entre os participantes, o que indicaria semelhanças entre os sujeitos trans (HT e MT) e os sujeitos cis (HC e MC) na manifestação de proeminências por meio do parâmetro da intensidade. Vale ressaltar, porém, as exceções observadas em HT2 e MT3, que apresentaram medidas discrepantes em relação aos demais sujeitos.

Com os resultados descritivos apresentados, esboçamos conclusões que nos permitem melhor abordar do ponto de vista linguístico o estilo de fala de sujeitos trans influenciadores digitais brasileiros conforme apresentado nas mídias digitais. Essas conclusões são apresentadas a seguir, com as quais finalizamos o presente artigo.

5 Conclusão

Este artigo explorou parâmetros da qualidade vocal e de recursos prosódicos em amostras de fala de influenciadores digitais transgênero (HT e MT) e cisgênero (HC e MC) a fim de atender dois objetivos: (i) comparar a julgamento perceptivo-auditivo atribuído à fala de sujeitos transgênero influenciadores digitais brasileiros com os parâmetros acústicos de natureza vocal e prosódica que caracterizam a fala desse sujeitos; (ii) comparar os parâmetros acústicos que caracterizam prosodicamente a fala de sujeitos transgênero em relação à fala de sujeitos cisgênero.

Descritivamente, a comparação entre os resultados encontrados na avaliação perceptivo-auditiva sobre o gênero do falante e os da análise acústica da voz pela medida de f_0 apresentou correspondência na maioria das vozes analisadas, mas não em todas. As vozes de HT apresentaram a medida de f_0 dentro da faixa de frequência esperada para o gênero masculino, entretanto, dois HT não foram identificados com o gênero declarado. Em relação às MT, a única voz que foi julgada como masculina apresentou a medida de f_0 na faixa de frequência esperada para o gênero masculino. Entretanto, vale ressaltar que uma segunda voz, na mesma faixa de frequência, a masculina, foi considerada feminina. Desses resultados, a interpretação que fizemos dos dados nos faz propor que o julgamento da manifestação vocal dos sujeitos trans não se sustenta pela recuperação de um padrão acústico; ao contrário, o julgamento é simbólico e parece congregar diferentes informações, tais como padrões de natureza prosódica, além de outros apontados pela literatura inclusive de natureza social, mas que não foram abordados pelo presente estudo. A esse respeito, a interpretação que fizemos desses resultados alinham-se com a proposição de Barros Filho (2005) de que a expressão vocal é parte da significação do enunciado e trata-se de uma expressão socialmente construída e não baseada em critérios fisiológicos, marcados especialmente nos parâmetros acústicos.

No que diz respeito aos recursos prosódicos conforme eles foram julgados na avaliação perceptivo-auditiva, fronteiras e proeminências foram identificadas pelos juízes experientes. No caso das fronteiras, embora possa haver variação entre sujeitos, observou-se que, na fala de HT e MT, as fronteiras identificadas foram em termos de frequência similares às observadas em HC e MC, sugerindo, assim, que o estilo de fala dos sujeitos trans no que diz respeito às proeminências foi avaliado como congruente ao gênero declarado. Interessantemente, na análise acústica, observou-se que o grupo HT utilizou uma faixa de frequência mais alta que o HC, marcando fronteiras com eventos tonais de forma mais aguda. Em contrapartida, o grupo MT utilizou uma faixa de frequência mais baixa que MC, marcando fronteiras de forma mais grave. Tem-se, portanto, que, se por um lado, do ponto de vista da avaliação perceptivo-auditiva, a frequência de proeminências não diferenciou sujeitos trans e cis; por outro lado, do ponto de vista acústico, diferenças observadas nos valores de f_0 que caracterizam as fronteiras prosódicas apontam para idiossincrasias próprias do grupo trans, o que nos sugere que esse seja um dos fatores que contribuem para a identificação de um estilo de fala próprio do grupo de sujeitos trans, relevante para se considerar uma identidade de grupo em uma nova representação identitária emergente na sociedade brasileira e marcada nas redes sociais: a dos sujeitos transgêneros.

O uso de marcação de proeminência se mostrou também como um fator que contribui para a identificação de um estilo de fala dos sujeitos trans, uma vez que, na avaliação perceptivo-auditiva, o julgamento das proeminências apontou que tanto HT quanto MT exploraram

menos esse recurso, em termos de frequência, do que os sujeitos cis (HC e MC), sugerindo, assim, que o estilo de fala dos sujeitos trans se distancia do estilo de fala do gênero declarado. Somam-se ainda duas outras observações advindas da análise acústica que corroboram a interpretação de um estilo de fala dos sujeitos trans: (a) no grupo HT, observou-se que a diferença entre valor mínimo e máximo de f_0 foi menor do que no HC, indicando uma tendência do grupo HT a uma fala menos marcada em relação a HC na sinalização das proeminências; (b) no grupo MT, a variação de f_0 se mostrou numa faixa de frequência mais baixa em relação à faixa usada por MC, indicando proeminências marcadas de forma mais grave.

Desses resultados, destacamos que tanto a avaliação perceptivo-auditiva quanto a análise acústica mostram a existência de proeminências menos frequentes e menos marcadas nos grupos trans quando comparados aos sujeitos cis; esses são dados que refutam a imagem de exagero, de caricatura e de estranhamento associada à fala trans pelo senso comum, usadas como forma de justificar uma suposta “anormalidade” considerada quando se assume a normatividade binária como padrão. Sugere-se, pois, nesse estudo, que esses juízos de valor partem de pré-construídos que não encontram respaldo na caracterização prosódica da fala desses sujeitos.

Vale retomar ainda que, no que diz respeito às fronteiras, a análise acústica mostrou também que as fronteiras identificadas na avaliação perceptivo-auditiva nem sempre foram marcadas pela produção de pausa e, quando o foram, houve grande variação na duração das pausas na fala de todos os participantes, fato que não permitiu identificar diferenças entre grupos, além de ser necessário considerar que tal variação pode estar associada à edição dos vídeos para postagem nas mídias sociais. Análise da intensidade na marcação de proeminências foi também outro parâmetro no qual se observou mais semelhanças entre os sujeitos trans (HT e MT) e os sujeitos cis (HC e MC).

Do conjunto de resultados apresentados, esboçamos duas conclusões gerais, que respondem assim aos objetivos delineados:

- 1 no cotejamento da avaliação perceptivo-auditiva e da análise acústica, identificam-se pontos de congruência e incongruência entre elas; os pontos de incongruência, especialmente, nos permitem concluir que a fala dos sujeitos trans, seja por sua expressão vocal, seja por suas características prosódicas, não é julgada exclusivamente por meio da recuperação de padrões acústicos; ao contrário, o julgamento perceptivo-auditivo é simbólico na medida em que algumas medidas acústicas não amparam o julgamento que se dá à fala dos sujeitos trans, seja pela categorização das vozes como masculina, feminina ou indefinida, seja pelas características atribuídas à fala desses sujeitos como “desviantes” ou “não naturais”; esses julgamentos, portanto, quando acontecem, existem *a priori* e são socialmente construídos.
- 2 a comparação dos parâmetros acústicos mensurados na fala de sujeitos trans e sujeitos cis mostra que a fala de sujeitos trans é dialeticamente constituída, pois existem pontos de aproximação e distanciamento entre a caracterização de aspectos vocais e prosódicos da fala desses sujeitos e a caracterização desses mesmos aspectos na fala do gênero declarado. Se, por um lado, há pontos de aproximação, há, por outro, pontos de distanciamento que nos levam a considerar a emergência de um estilo de fala dos sujeitos trans como resultante do deslocamento desses sujeitos frente ao vazio simbólico identificado por esses sujeitos na relação com os outros. Tal estilo

é interpretado aqui como uma nova configuração de representações identitárias – explicada pelo rompimento com o padrão binário normatizante – que se mostra no discurso produzido e disseminado nas plataformas digitais, inclusive plasmado nos recursos prosódicos. Cabe ressaltar que a população estudada neste trabalho foi composta por sujeitos transgêneros influenciadores digitais, o que implica o perfil de pessoas que têm uma prática comunicativa não amadora no ambiente digital. Seria interessante que outros estudos fossem desenvolvidos com sujeitos transgênero não envolvidos na mídia social como influenciadores para avaliar se os resultados aqui apresentados se estenderiam também a eles.

Pesquisas futuras também poderiam aprofundar e ampliar o conhecimento sobre o estilo de fala dos sujeitos trans com amostras maiores, já que o número restrito de sujeitos cujas amostras de fala foram analisadas é uma limitação do estudo, o que impossibilitou a realização de análise estatística inferencial de modo a validar estatisticamente o que observamos descritivamente e, assim, nos permitir fazer afirmações de forma mais robusta. Destaca-se, ainda, que análises com a inserção de novos parâmetros, como velocidade de fala, articulação, entre outros, são bem-vindas para melhor entendimento do tema.

Declaração de autoria

Geovana Soncin e Eryne Alves Bafum participaram da elaboração do projeto de pesquisa, levantamento bibliográfico, desenvolvimento da pesquisa, discussão dos dados e escrita do artigo. Gabriela Aparecida Rodrigues Gonçalves, Giovanna Caroline Borges e Karoline Araujo dos Santos participaram do desenvolvimento da pesquisa, realizando parte dos procedimentos metodológicos e participaram da discussão dos dados analisados. Eliana Maria Gradim Fabbron supervisionou a elaboração do projeto de pesquisa e participou do levantamento bibliográfico, do desenvolvimento da pesquisa, da discussão dos dados e da organização da escrita do artigo. As primeira, segunda e quinta autoras respondem pelas decisões tomadas para o texto após avaliação dos pareceristas.

Agradecimentos

Agradecemos à FAPESP pelo financiamento à pesquisa para a primeira, terceira e quinta autora no âmbito do projeto 2020/10144-3 e aos pareceristas que avaliaram o trabalho que, com suas sugestões teóricas e metodológicas, contribuíram fortemente para o enriquecimento do texto.

Referências

- ALVES, S. Brasil é o 3º país do mundo que mais buscou por LGBTQIA+ no último ano, segundo dados do Google. Disponível em: <https://epocanegocios.globo.com/Tecnologia/noticia/2022/06/brasil-e-o-3-pais-do-mundo-que-mais-buscou-por-lgbtqia-no-ultimo-ano-segundo-dados-do-google.html>. 28 Jun 2022 – 18h27. Atualizado em 28 Jun 2022. Acesso em: dezembro de 2023.
- ASTÉSANO, C. et al. Prosodic cues to syntactic structure in French: The role of prosodic boundaries. *Journal of Memory and Language*, Amsterdã, v. 51, n. 1, p. 95-109, 2004. DOI: <https://doi.org/10.1016/j.jml.2004.02.003>.
- BAKHTIN, M. Estética da criação verbal. São Paulo: Martins Fontes, 2003.
- BAKHTIN, M. *Problemas da poética de Dostoievski*. Trad. Paulo Bezerra. 2 ed. Rio de Janeiro: Forense Universitária, 1981.
- BAKHTIN, M.; VOLOCHINOV, V. Discurso na vida e discurso na arte. Tradução de Cristóvão Tezza. In: Voloshinov, V.N. *Freudism: a marxist critique*. New York: Academic Press, 1976. p. 93-116.
- BAKHTIN, M.; VOLOCHINOV, V. *Marxismo e filosofia da linguagem*. São Paulo: Hucitec, 2004.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas: Pontes, 2006.
- BARBOSA, P. A.; MADUREIRA, S. *Manual de Fonética Acústica Experimental: aplicações a dados do Português*. São Paulo: Cortez, 2015.
- BARROS FILHO, C. A construção social da Voz. In: KYRRILLOS, L. R. *Expressividade: da teoria à prática*. Rio de Janeiro: Revinter, 2005. p. 45-60. DOI: <https://doi.org/10.15448/1980-3729.2004.23.3253>.
- BENEVIDES, B. G. *Dossiê: assassinatos e violências contra travestis e transexuais brasileiras em 2022*. Brasília, DF: Distrito Drag; ANTRA (Associação Nacional de Travestis e Transexuais), 2023. 109 p.
- BOERSMA, P.; WEENINK, D. Praat. Doing phonetics by computer (Version 5.1), 2005. Disponível em: <https://www.fon.hum.uva.nl/praat/>. Acesso em: 10 de abril de 2024
- BREEN, M.; et al. Acoustic correlates of information structure. *Language and Cognitive Processes*, Londres, v. 25, n. 7-9, p. 1044–1098, 2010. DOI: <https://doi.org/10.1080/01690961003734037>.
- CAGLIARI, L. C. Prosódia: algumas funções dos supra-segmentos. *Cadernos de estudos linguísticos*, Campinas, v. 23, p. 137-151, 1992. DOI: <https://doi.org/10.20396/cel.v23i0.8636850>
- CANAL, M. F. et al. Perceptual-auditory and acoustic analysis of breathiness in cis and transgender men and women. *Journal of Voice*, Philadelphia, no prelo, 2024. DOI: <https://doi.org/10.1016/j.jvoice.2023.06.020>.
- CARPES, D. F. R. P. *Comportamento prosódico e sintático do foco em PB: um estudo experimental de interface*. 2019. 150f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2019.
- CUTLER, A.; NORRIS, D. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, Washington, v. 14, n. 1, p. 113–121, 1988. DOI: <https://doi.org/10.1037/0096-1523.14.1.113>.

DAHL, K. L.; MAHLER, L. A. Acoustic features of transfeminine voices and perceptions of voice femininity. *Journal of Voice*, Philadelphia, v. 34, n. 6, p. 961.e19-961.e26, 2020. DOI: <https://doi.org/10.1016/j.jvoice.2019.11.002>.

DIAS, E. C. A diferença entre transexual e transgênero: de que se trata para a psicanálise? *Correio Express – Revista Online Escola Brasileira de Psicanálise*. Disponível em: https://www.ebp.org.br/correio_express/2021/07/29/a-diferenca-entre-transexual-e-transgenero-de-que-se-trata-para-a-psicanalise/#_edn5. 29 de julho 2021. Acesso em: setembro de 2024.

D'IMPERIO, M., et al. Intonational Phrasing in Romance: The role of prosodic and syntactic structure. In: FROTA, S., VIGÁRIO, M.; FREITAS, M. J. (eds.). *Prosodies: with special reference to Iberian Languages*. Phonetics & Phonology Series. Berlin: Mouton de Gruyter, 2005, p. 59-97. DOI: <https://doi.org/10.1515/9783110197587>.

DRUMMOND, L. B. Fonoaudiologia e transgenitalização: a voz no processo de reelaboração da identidade social do transexual. In: ASSOCIAÇÃO BRASILEIRA DE PSICOLOGIA SOCIAL. XV Encontro Nacional da Associação Brasileira de Psicologia Social, Belo Horizonte, 2009. *Anais...* Belo Horizonte: Associação Brasileira de Psicologia Social, 2009. p. 150-165.

ERICSDOTTER C, ERICSSON AM. Gender differences in vowel duration in read Swedish: Preliminary results. *Working Papers*, Stockholm, v. 49, p. 34–37, 2001.

FOWLER, C. A. Listeners do hear sounds, not tongues. *The Journal of the Acoustical Society of America*, New York, v. 99, n. 3, p. 1730-1741, 1996. DOI: <https://doi.org/10.1121/1.414692>.

GELFER, M. P.; BENNETT, Q. E. Speaking fundamental frequency and vowel formant frequencies: effects on perception of gender. *Journal of Voice*, Philadelphia, v. 27, n. 5, p. 556-566, 2013. DOI: <https://doi.org/10.1016/j.jvoice.2013.03.007>.

GELFER, M. P.; SCHOFIELD, K. J. Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *Journal of Voice*, Philadelphia, v. 14, n.1, p. 22-33, 2000. DOI: [https://doi.org/10.1016/S0892-1997\(00\)80092-2](https://doi.org/10.1016/S0892-1997(00)80092-2).

GOLDSTEIN, L.; FOWLER, C. A. Articulatory phonology: A phonology for public language use. In: BEAUCHAMP, D.; FRANÇOISE, H.; LECOURS, M. (Org.) *Phonetics and phonology in language comprehension and production: Differences and similarities*. [S.I.]: Walter de Gruyter, 2003. p. 159-207. DOI: <https://doi.org/10.1515/9783110895094>.

GORHAM-ROWAN, M.; MORRIS, R. Aerodynamic analysis of male-to-female transgender voice. *Journal of Voice*, Philadelphia, v. 20, n. 2, p. 251-262, 2006. DOI: <https://doi.org/10.1016/j.jvoice.2005.03.001>.

GOVERNO DO ESTADO DE SÃO PAULO. COORDENADORIA DE GESTÃO DA EDUCAÇÃO BÁSICA. DOCUMENTO ORIENTADOR CGEB: TRATAMENTO NOMINAL DE DISCENTES TRAVESTIS E TRANSEXUAIS. São Paulo, 2015. Disponível em: <https://midiastoragesec.blob.core.windows.net/001/2017/10/2016-14-07-nomesocial.pdf>. Acesso em: 9 jul. 2022.

GUSSENHOVEN, C. Types of focus in English. In: LEE, C.; GORDON, M.; BURING, D. (eds.). *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation*. Dordrecht: Springer, 2006, p. 83-100. DOI: <https://doi.org/10.1007/978-1-4020-4796-1>.

HARDY, T. L. D. et al. Acoustic Predictors of Gender Attribution, Masculinity. *Journal of Voice*, Philadelphia, v. 34, n. 2, p. 300.e11-300.e26, 2018. DOI: <https://doi.org/10.1016/j.jvoice.2018.08.005>.

- HILLENBRAND, J. et al. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, New York, v. 97, n. 5, p. 3099-3111, 1995. DOI: <https://doi.org/10.1121/1.411872>.
- HILLENBRAND, J. M.; CLARK, M. J. The role of fo and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics*, Berlim, v. 71, n. 5, p. 1150-1166, 2009. DOI: <https://doi.org/10.3758/APP.71.5.1150>.
- HIRST, D.; DI CRISTO, A. *Intonation systems: A Survey of Twenty Languages*. Cambridge: CUP, 1998
- HOULE, N.; LEVI, S. V. Effect of phonation on perception of femininity/masculinity in transgender and cisgender speakers. *Journal of Voice*, Philadelphia, v. 35, n. 3, p. 497. e23-497. e37, 2021. DOI: <https://doi.org/10.1016/j.jvoice.2020.07.013>.
- JOHNSON, K.; MARTIN, J. Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica*, Basel, v. 58, n. 1-2, p. 81-102, 2001. DOI: <https://doi.org/10.1159/000028489>
- LADD, D. R. *Intonational phonology*. Cambridge: University Press, 1996
- LEHISTE, I. The perception of duration within sequences of four intervals. *Journal of Phonetics*, Amsterdā, v. 7, n. 4, p. 313-316, 1979.
- LEUNG, Y.; OATES, J.; CHAN, S. P. Voice, articulation, and prosody contribute to listener perceptions of speaker gender: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, Rockville, v. 61, n. 2, p. 266-297, 2018. DOI: https://doi.org/10.1044/2017_JSLHR-S-17-0067.
- LEVELT, W. J. M. *Speaking: From intention to articulation*. Cambridge: MIT Press, 1993.
- LIBERMAN, A. M.; MATTINGLY, I. G. The motor theory of speech perception revised. *Cognition*, Amsterdā, v. 21, n. 1, p. 1-36, 1985. DOI: [https://doi.org/10.1016/0010-0277\(85\)90021-6](https://doi.org/10.1016/0010-0277(85)90021-6).
- LIMA, A. M.; CONSTANTINI, A. C. Prosódia e fonoaudiologia: do fonoestilo ao transtorno da linguagem. In: FREITAG, R. M. K.; LUCENTE, L. (orgs.) *Prosódia da fala: pesquisa e ensino*. São Paulo: Blucher, 2017. p. 133 -144.
- MATOS, M. Teorias de gênero ou teorias e gênero? Se e como os estudos de gênero e feministas se transformaram em um campo novo para as ciências. *Revista Estudos Feministas*, Florianópolis, v. 16, n. 2, p. 333-357, 2008. DOI <https://doi.org/10.1590/S0104-026X2008000200003>.
- MARKOVÁ, I. *Dialogicidade e Representações Sociais: As dinâmicas da mente*. Petrópolis, RJ: Vozes, 2006.
- MCNEILL, E. J. M; et al. Perception of Voice in the Transgender Client. *Journal of Voice*, Philadelphia, v. 22, n. 6, p. 727-733, 2006. DOI: <https://doi.org/10.1016/j.jvoice.2006.12.010>.
- MERLO, S.; BARBOSA, P. A. Séries temporais de pausas e de hesitações na fala espontânea. *Cadernos de Estudos Linguísticos*, Campinas v. 54, n. 1, p. 11-24, 2012.
- MINISTÉRIO DOS DIREITOS HUMANOS (Brasil) (org.). Manual orientador sobre diversidade. Disponível em: https://www.gov.br/mdh/ptbr/assuntos/noticias/2018/dezembro/ministerio-lanca-manual-orientador-dediversidade/copy_of_ManualLGBTDIGITAL.pdf. Acesso em: 29 jul. 20
- MOSCOVICI, S. *La psychanalyse, son image et son public*. Paris: Press Universitaires de France, 1961.
- MUNSON, B. The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech*, Thousand Oaks, v. 50, p. 125-142, 2007. DOI: <https://doi.org/10.1177/00238309070500010601>.

NASCIMENTO, J. C.; CHACON, L. Por uma visão discursiva do fenômeno da hesitação. *ALFA: Revista de Linguística*, São Paulo, v. 50, n. 1, p. 59-76, 2006. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/1395>.

NEUMANN, K.; WELZEL, C. The importance of the voice in male-to-female transsexualism. *Journal of Voice*, Philadelphia, v. 18, n. 1, p. 153-167, 2004. DOI: [https://doi.org/10.1016/S0892-1997\(03\)00084-5](https://doi.org/10.1016/S0892-1997(03)00084-5).

NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.

OWEN, K.; HANCOCK, A. B. The role of self- and listener perceptions of femininity in voice therapy. *International Journal of Transgenderism*, v. 12, n. 4, p. 272-284, 2010. DOI: <https://doi.org/10.1080/15532739.2010.550767>.

PAPELEU, T. et al. Intonation parameters in gender diverse people. *Journal of Voice*, Philadelphia, [s.l.], [no prelo], 2023. DOI: <https://doi.org/10.1016/j.jvoice.2022.12.020>.

PATTERSON, R. D. Perception of prosody. In: REMEZ, R.; IVERSON, P. B. (eds.). *The Handbook of Speech Perception*. Wiley-Blackwell, 2019. p. 311-340.

PIJPER, J. R.; SANDERMAN, A. A. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America*, New York, v. 96, n. 4, p. 2037-2047, 1994. DOI: <https://doi.org/10.1121/1.410145>.

PINTO, J. P. Pragmática. In: MUSSALIM, F.; BENTES, A. C. (orgs.). *Introdução à Linguística: domínios e fronteiras*, volume 2. 8ª ed. São Paulo: Cortez, 2012.

RIBEIRO, P. B.; SOBRAL, A. Eu, o outro (Outro) e o vazio na constituição da representação identitária. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 37, n. 1, 2021, p. 1-25. DOI: <https://doi.org/10.1590/1678-460X2021370110>.

REIS, T., org. *Manual de Comunicação LGBTI+*. 2ª edição. Curitiba: Aliança Nacional LGBTI / GayLatino, 2018.

SANTOS, K. A. dos, et al. Focalização prosódica na fala de crianças com desenvolvimento fonológico típico e atípico: análise duracional. *Veredas Revista de Estudos Linguísticos*, Juiz de Fora, v. 21, n. 1, p. e40928, 2023. DOI: <https://doi.org/10.34019/1982-2243.2023.v27.40928>.

SCHMIDT, J. G. et al. O desafio da voz na mulher transgênero: autopercepção de desvantagem vocal em mulheres trans em comparação à percepção de gênero por ouvintes leigos. *Revista CEFAC*, São Paulo, v. 20, p. 79-86, 2018. DOI: <https://doi.org/10.1590/1982-021620182011217>.

SERRA, C. R. *Realização e percepção de fronteiras prosódicas no português do Brasil: fala espontânea e leitura*. 2009. 214 f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio Janeiro, Rio de Janeiro, 2009.

Silva, E. R.; Oliveira, S. M. D. A.; Silva, M. G. P. Promoção à saúde vocal em homens transgêneros. *Distúrbios da Comunicação*, 33(1), 173-177, 2021.

SONCIN, G. Alongamento em fronteira de frase entoacional no Português do Brasil: evidências a partir de um design experimental. *GRADUS: Revista Brasileira de Fonologia de Laboratório*, Porto Alegre, v. 3, n. 1, p. 64-80, 2018. DOI: <http://dx.doi.org/10.47627/gradus.v3i1.119>

SONCIN, G.; TENANI, L. E. Variações de F_0 e configurações de frase entoacional: análise de estruturas contrastivas. *Domínios de Lingu@gem*, Uberlândia, v. 10, n. 2, p. 534-558, 2016. DOI: [10.14393/DL22-v10n2a2016-6](https://doi.org/10.14393/DL22-v10n2a2016-6).

- SONCIN, G.; TENANI, L.; BERTI, L. Percepção de pausa em fronteira prosódica. *SCRIPTA*, Belo Horizonte, v. 4, n. 21, p. 143-164, 2017. DOI: <https://doi.org/10.5752/P2358-3428.2017V21n41p143>
- SONCIN, G.; TENANI, L.; BERTI, L. Phonologic representation and speech perception: the role of pause. *Diacritica*, Minho, v. 33, n. 2, p. 4-18, 2019. DOI: doi.org/10.21814/diacritica.401
- SPAZZAPAN, E. A. et al. Características acústicas da voz em diferentes ciclos da vida: revisão integrativa da literatura. *Revista CEFAC*, São Paulo, v. 21, p. e15018, 2019. DOI: <https://doi.org/10.1590/1982-0216/201921315018>.
- SPAZZAPAN, E. A. et al. Smoothed cepstral peak analysis of Brazilian children and adolescents speakers. *Journal of Voice*, Philadelphia, v. 38, n.5, p. 1149–1155, 2022. DOI: <https://doi.org/10.1016/j.jvoice.2022.02.002>.
- SPIZZIRRI, G. et al. Proportion of ALGBT adult Brazilians, sociodemographic characteristics, and self-reported violence. *Scientific Reports*, Londres, v. 12, n. 1, p. 11176, 2022. DOI: <https://doi.org/10.1038/s41598-022-15103-y>.
- SWERTS, M.; GELUYKENS, R. Prosody as a marker of information flow in spoken discourse. *Language and speech*, Thousand Oaks, v. 37, n. 1, p. 21-43, 1994. DOI: <https://doi.org/10.1177/002383099403700102>.
- TERKEN, J.; HERMES, D. The perception of prosodic prominence. In: HORNE, M. (ed.). *Prosody: Theory and Experiment*. Studies presented to Gösta Bruce. Dordrecht: Springer Netherlands, 2000. p. 89-127. DOI: https://doi.org/10.1007/978-94-015-9413-4_5.
- VAN BORSEL, J.; DE POT, K.; DE CUYPERE, G. Voice and physical appearance in female to-male transsexuals. *Journal of Voice*, Philadelphia, v. 23, n. 4, p. 494-497, 2009.
- VIEIRA, V. F. *Padrão entoacional e duracional da fala de mulheres transexuais*. 2018. 28 f. Trabalho de Conclusão de Curso (Bacharelado em Fonoaudiologia) – Universidade de Brasília, Brasília, 2018.
- VILLAS-BÔAS, A. P. et al. Acoustic Measures of Brazilian Transgender Women's: A Case-Control Study. *Frontiers in Psychology*, Lausanne, v. 12, n. 12, p. 622526, 2021. DOI: <https://doi.org/10.3389/fpsyg.2021.622526>.
- ZELLNER, B.. Pauses and the temporal structure of speech. In: KELLER, E. (Ed.). *Fundamentals of speech synthesis and speech recognition*. Chichester: John Wiley, 1994. p. 41-62.
- WOLFE, V.I.; RATUSNIK, D.L.; SMITH, F.H.; NORTHROP, G. Intonation and Fundamental Frequency in Male-to-Female Transsexuals. *Journal of Speech and Hearing Disorders*, Rockville, v. 55, n. 1, p. 43- 50, 1990. DOI: <https://doi.org/10.1044/jshd.5501.43>.

O redobro do sujeito ele tem duas realizações prosódicas: o fraseamento prosódico das construções de redobro do sujeito na fala do Rio de Janeiro

Subject Doubling it has two Prosodic Realizations: Prosodic Phrasing of the Subject Doubling Constructions in the Speech of Rio de Janeiro

Eduardo Patrick Rezende dos Reis
Universidade Federal do Rio de Janeiro
(UFRJ) | Rio de Janeiro | RJ | BR
eduardorezende@letras.ufrj.br
<https://orcid.org/0000-0002-5049-4200>

Resumo: Este trabalho tem como objetivo investigar a(s) realização(ões) prosódica(s) das construções de redobro do sujeito em dados de falantes do Rio de Janeiro, visando a demonstrar que, na verdade, as estruturas sob esse rótulo não apresentam comportamento acústico / entoacional uniforme (cf. Rezende Dos Reis, 2023). A amostra foi extraída do *Corpus Concordância*, que compõe o Projeto COMPARAPORT, bem como da plataforma *Youtube* e do canal midiático *Globo News*. Para o tratamento prosódico, foi utilizado o *software* de análise acústica PRAAT (Boersma; Weenink, 2023); o tratamento estatístico, por sua vez, foi realizado com o auxílio do programa MINITAB versão 21.1. Partindo sobretudo de Rezende dos Reis (2023), a hipótese levantada é a de que há, no Português Brasileiro, duas estruturas prosódicas para o redobro do sujeito, fraseadas diferentemente. Amparado pela hierarquia Prosódica (cf. Nespor; Vogel, 2007 [1986]) e pela Fonologia Entoacional (cf. Ladd, 2008 [1996]), este trabalho argumenta que, quando o DP inicial estiver deslocado (constituindo um verdadeiro “tópico”), ele será fraseado em um sintagma entoacional (*Intonational Phrase* - IP) independente da sentença-comentário, conforme já apontado em trabalhos anteriores (cf. Silva, 2018; Yano; Fernandes, 2020); somados a essas ocorrências, ainda podem ser atestados casos em que o DP inicial e a sentença que o segue são mapeados em um único IP. A partir desse cenário, este trabalho propõe ainda que



a verificação da ocorrência de fronteira prosódica configura um eficiente critério atuante na diferenciação das realizações prosódicas do redobro do sujeito. Os resultados obtidos confirmam a referida hipótese.

Palavras-chave: o redobro do sujeito no PB; fonologia prosódica; fonologia entoacional; fraseamento prosódico.

Abstract: This paper investigates the prosodic realization(s) of subject doubling constructions in the speech data of Rio de Janeiro speakers, aiming to demonstrate that, in fact, the structures under this label do not exhibit uniform acoustic / intonational behavior (cf. Rezende Dos Reis, 2023). The data come from the *Corpus Concordância*, which is part of the “Projeto COMPARAPORT”, as well as from the *YouTube* platform and the media channel *Globo News*. For acoustic analysis, the computational program PRAAT was used (Boersma; Weenink, 2023); statistical analysis, in turn, was performed with the assistance of the MINITAB software version 21.1. Based specially on Rezende dos Reis (2023), the hypothesis raised is that Brazilian Portuguese has two prosodic structures for subject doubling, which are phrased differently. Supported by the Prosodic Hierarchy (cf. Nespor; Vogel, 2007 [1986]) and the Intonational Phonology (cf. Ladd, 2008 [1996]), this paper argues that, when the initial DP is displaced (thus constituting a true “topic”), it will be phrased as an independent intonational phrase (IP) from the sentence-comment, as already pointed out in previous works (cf. Silva, 2018; Yano; Fernandes, 2020); in addition to these occurrences, there are cases where the initial DP is mapped along with the sentence following it in a single IP. Based on this scenario, this work also proposes that the verification of the occurrence of prosodic boundaries constitutes an efficient criterion to differentiate the prosodic realizations of subject doubling. The obtained results confirm the aforementioned hypothesis.

Keywords: subject doubling in BP; prosodic phonology; intonational phonology; prosodic phrasing.

1 Pontos de partida

Desde a década de 1980, uma gama de estudos tem se debruçado, em especial no contexto das línguas românicas, sobre as configurações sintáticas denominadas de “construções de tópico marcado” (CTs) (cf. Brito; Duarte; Matos, 2003). Focalizando especificamente o Português do Brasil (PB), não se pode deixar de pontuar os trabalhos pioneiros de Eunice Pontes, compilados no livro *O tópico no Português do Brasil* (Pontes, 1987). Recorrentemente inscrita no rol das CTs, encontra-se aquela a que Pontes chama de “Deslocamento à Esquerda” de sujeito (DE), que será referenciada neste trabalho como “redobro do sujeito”, por razões a serem delineadas no decorrer desta seção. Na referida estrutura, visualiza-se a vinculação entre um DP (*Determiner Phrase*) em posição inicial da sentença e um pronome resumptivo na função de sujeito sintático (Duarte; Kato, 2008; Rezende Dos Reis, 2023), como se vê nos exemplos a seguir.

- (1) a. [O meu pai]₁ ele₁ estudou em escolas públicas a vida toda. (DP lexical)
b. *Eles*, aparentemente *eles*, tavam primeiro fazendo bagunça. (DP pronominal)

Segundo Pontes (1987), as construções de “Deslocamento à Esquerda” de sujeito estão entre as mais frequentes CTs, o que viria a ser confirmado em trabalhos posteriores (cf. Orsini, 2003; Vasco, 2006; Orsini; Vasco, 2007). Para Duarte (1995, 2003; entre outros), o redobro do sujeito no PB pode ser entendido, nos moldes de Weinreich, Labov e Herzog (1968), como um efeito colateral da remarcação no valor do Parâmetro do Sujeito Nulo (cf. Chomsky, 1981), de positivo para negativo. Ancorados em Duarte (1995), diversos trabalhos empíricos posteriores para a variedade brasileira têm evidenciado que as estruturas aqui examinadas parecem estar cada vez mais implementadas em tal sistema (cf. Orsini, 2003; Nicolau De Paula, 2012; Rezende Dos Reis, 2023; entre outros), em virtude das poucas restrições que atuam sobre o DP inicial que as compõe, bem como do contínuo aumento de sua ocorrência na fala (semi)espontânea.

Atualmente, entretanto, muito se tem discutido quanto à verdadeira natureza das construções de redobro do sujeito. Na compreensão de Pontes (1987), que se pauta em Ross (1967), o DP inicial nas referidas estruturas corresponde a um constituinte deslocado para a borda esquerda da sentença, podendo ser eventualmente seguido ou não de pausa silenciosa. Respaldados em Pontes, estudos subsequentes, que focalizam o PB e outras variedades do português (cf. Vasco, 2006), igualmente assumiram que tal DP se localiza obrigatoriamente no domínio de CP (*Complementizer Phrase*) (cf. Rizzi, 1997). O termo “Deslocamento à Esquerda” ganha “visibilidade” e começa a ser amplamente utilizado na literatura especializada (cf. Orsini, 2003; Orsini; Vasco, 2007; Nicolau De Paula, 2012; entre outros). Entretanto, tem-se verificado que as configurações sintáticas abrigadas por esse rótulo não exibem um comportamento uniforme, observação esta que tem suscitado uma série de questionamentos sobre a sua natureza, conforme se vê na sequência.

A partir dos anos 2000, ganham espaço trabalhos sobre o tema que adotam um caminho parcialmente distinto, segundo o qual as construções até então referidas como DEs de sujeito nem sempre devem ser interpretadas como uma estratégia de periferização do DP alvo do redobro; nem todo DP linearmente à esquerda, portanto, estaria efetivamente deslocado - isto é, em uma posição A' (cf. Costa; Duarte; Silva, 2004). Seguindo essa linha de

raciocínio, temos o trabalho de orientação gerativista de Gasque de Souza (2021), que visou a analisar, à luz de parâmetros acústicos e sintático-discursivos, a estratégia de redobro do sujeito em 17 inquéritos do PB, extraídos do *Corpus LínguaPOA*. Na referida análise, a autora detectou uma possível correlação entre propriedades do nível semântico-discursivo e do nível prosódico, o estatuto informacional do DP inicial redobrado e a ocorrência de “pausa”: os DPs que codificam informação “velha” tendem a ser seguidos de pausa; os que articulam informação nova, não. Com base nesses parâmetros, Gasque de Souza argumenta em favor de que as construções de redobro do sujeito sem pausa entre o DP e o pronome-cópia não equivalem a verdadeiras DEs, uma vez que o DP duplicado (presente nelas) não codificaria, obrigatoriamente, traços característicos da categoria “tópico”; não ocuparia, então, uma posição de tópico. As DEs “verdadeiras”, por outro lado, corresponderiam às estruturas em que se atesta a ocorrência de pausa entre o DP deslocado e a sentença-comentário.

Inscritos no programa de pesquisa “cartográfico” (cf. Cinque; Rizzi, 2010)¹, chamam a atenção os trabalhos de Quarezemin (2017, 2019; 2020) e, sobretudo, o de Krieck (2022), que igualmente defendem, a partir de um conjunto de evidências empíricas, que nem todo DP linearmente à esquerda está localizado no domínio de CP. Conforme se visualiza em (2), extraído de Krieck (2022, p. 78), parece haver restrições à ocorrência de redobro em contextos *out-of-the-blue*, quando constituído de um DP inicial indefinido.

(2) O que aconteceu?

- a. [Um carro], *ele*, bateu no poste.
- a.’ Pelo que eu supus, [um carro], *ele*, bateu no poste.
- ??? a.” [Um carro], pelo que eu supus, *ele*, bateu no poste.

Segundo Krieck (2022), apoiada em Rizzi (2005), a estranheza da construção (2a”) deriva da hipótese de que um DP indefinido, no contexto apresentado, não pode congelar em posição de tópico, visto que um DP tópico tende a evocar uma informação velha; o contexto *out-of-the-blue*, contudo, força uma estrutura que manifeste uma informação nova (foco largo). Em contrapartida, a gramaticalidade das estruturas em (2a-a’), atribuída ao não rompimento sintático entre o DP inicial e o pronome resumptivo (Krieck, 2022), permite conjecturar que aquele não se encontra em uma posição A’, mas em uma posição A, no domínio de TP (*Tense Phrase*). Uma outra evidência de que nem todo DP em estruturas de redobro congele na borda esquerda da sentença se verifica tendo em vista que, em tais construções, não se processa o efeito de minimalidade relativizada quando ocorre a extração de uma elemento *-wh* (3) (cf. Quarezemin, 2020; Quarezemin; Tescari Neto, 2024).

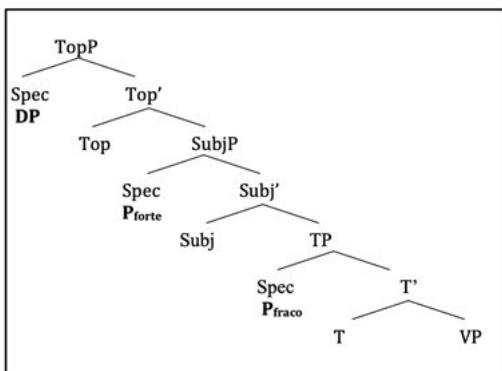
(3) a. Onde Pedro acha que a Ana ela encontrou João?

- a.’ *Onde Pedro acha que João a Ana (ela) encontrou?

¹ Segundo Tescari Neto (2022), o programa cartográfico consiste em uma vertente da Teoria de Princípios e Parâmetros, que visa à reflexão e à elaboração de “mapas” para as estruturas sintáticas das línguas naturais, investigando, detalhada e sistematicamente, a estrutura hierárquica que constitui a arquitetura da sentença. Tal abordagem busca, de modo objetivo, instituir uma relação simétrica entre propriedades morfossintáticas e semânticas e a composição do esqueleto estrutural da sentença.

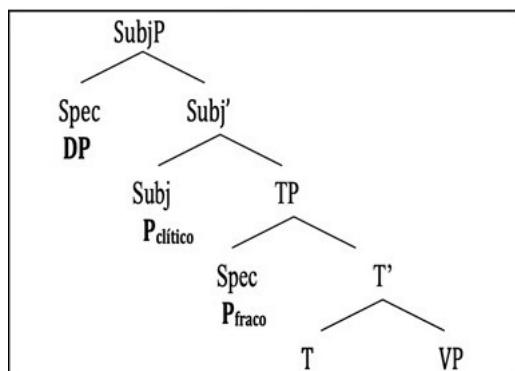
Nos exemplos em (3), a presença do DP “João”, mas não do DP redobrado “a Ana”, em posição “inicial” da sentença encaixada impede a extração do elemento-*wh* “Onde” desse domínio. Tanto Quarezemin (2017; 2019) quanto Krieck (2022) concluem, portanto, que o DP inicial em construção de redobro no PB, dado o seu comportamento sintático não uniforme, pode ocupar uma posição no domínio não argumental, Spec de TopP (cf. Rizzi, 1997), tal como uma posição hierarquicamente inferior, projetada no domínio argumental, Spec de SubjP²³(cf. Cardinaletti, 2004, 2014), conforme se percebe nas Figuras 1 e 2.

Figura 1 – Formalização do redobro do sujeito “verdeiro”



Fonte: Krieck (2022).

Figura 2 – Formalização do redobro so sujeito “falso”



Fonte: Krieck (2022).

Tal qual ressalta Rezende dos Reis (2023), a dissertação de Krieck (2022) fornece um conjunto de propriedades prosódico-sintáticas e distribucionais⁴ (resultado da aplicação de operações sintáticas) que parecem viabilizar a diagnose da posição ocupada pelo DP inicial. Segundo Krieck, o DP será analisado como um constituinte deslocado em caso de: (a) apresentar um contorno entoacional de tópico (cf. Rizzi, 1997),⁵

² O projeto de pesquisa em cartografia sintática adota, como um de seus pressupostos epistemológicos, a máxima *One Feature, One Head* (cf. Kayne, 2005; Tescari Neto, 2022), segundo a qual cada traço conceitual presente nos sistemas linguísticos projetará uma camada funcional. Neste aspecto, a camada funcional SubjP é projetada mediante a checagem do traço “sujeito da predicação” pelo “sujeito semântico” da derivação, que não necessariamente coincidirá com o “sujeito sintático”, aquele que checa os traços-*phi* e o caso nominativo e que congela em Spec de TP (cf. Cardinaletti, 2004; Krieck, 2022).

³ Uma alternativa ao tratamento cartográfico pode ser conferida em Rezende dos Reis e Duarte (no prelo), que propõem uma formalização para cada tipo de redobro, partindo do “modelo de herança de traços” de Miyagawa (2007; 2010). Em vez de considerar a dicotomia tópico (deslocado) x sujeito da predicação (não deslocado), os autores consideram que o redobro é sempre composto de um “DP tópico”, que pode congelar tanto em uma posição A', Spec de TopP, quanto em uma posição A, Spec de aP (cf. Miyagawa, 2010).

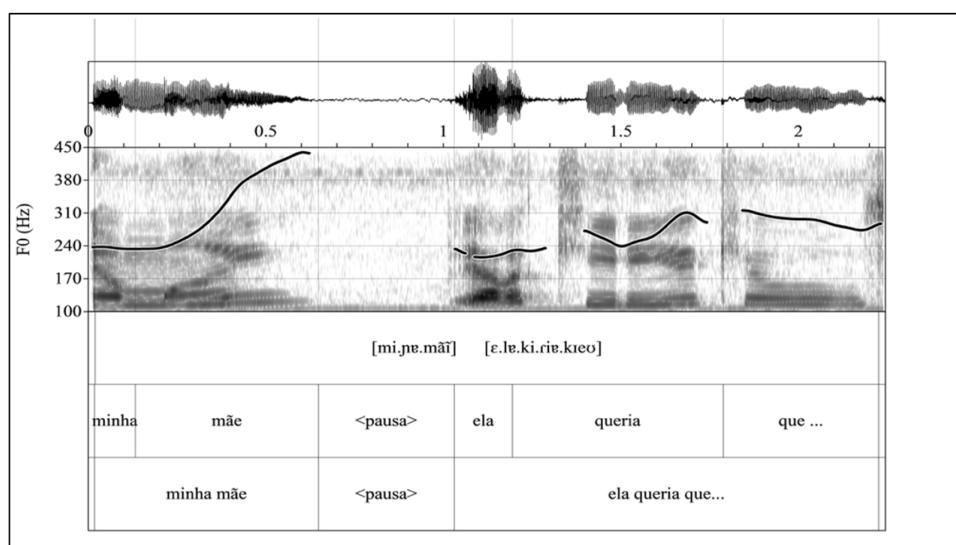
⁴ Enfatizo que o trabalho de Krieck (2022) se arquiteta numa abordagem cartográfica, que, embora estreite o diálogo com os níveis prosódico-discursivos, corresponde a uma perspectiva “sintaticocêntrica” (cf. Cinque; Rizzi, 2010): a partir da aplicação de operações sintáticas, o processo derivacional gerará um “material” que será lido pelos sistemas de interface “articulatório-perceptual” e “conceitual-intencional”. O conjunto de propriedades a que Krieck chega seria, portanto, um reflexo da atuação de um “dispositivo sintático intensional”. O trabalho de Krieck (2022), pois, não configurou um empreendimento com o comprometimento de realizar uma análise acústica dos dados de redobro do sujeito; todavia, a ausência dessa análise a impossibilitou de alcançar uma generalização, essencialmente quanto ao critério prosódico, mais precisa.

⁵ Segundo Rizzi (1997), referência na qual Krieck se apoia, o constituinte tópico é um elemento anteposto à sentença-comentário, da qual está desvinculado por uma “entonação da vírgula”. Tal definição se mostra, no

seguido ou não de pausa entre esse DP e o pronome nominativo; e/ou (b) apresentar um constituinte interpolado, responsável pela quebra da adjacência sintática entre os referidos expedientes sintáticos. Caso não se verifiquem as “condições” mencionadas, o DP redobrado, apesar de estar linearmente em posição inicial, ocuparia uma posição na zona argumental, hierarquicamente mais baixa que a de um DP tópico prototípico. Pressuposto a tal compreensão, está codificado o entendimento de que cada um dos tipos de redobro representaria um algoritmo derivacional específico, que refletiria traços prosódicos e semântico-discursivos singulares na “estrutura de superfície”. Por conseguinte, Quarezemin (2019) e Kriech (2022) alertam para o cuidado que se deve ter na utilização do rótulo “DE”, uma vez que nem tudo inscrito sob esse metatermo, ao contrário do que sugere, se encontraria efetivamente deslocado. Devido a isso, compreendo que o rótulo “redobro do sujeito” consiste em uma “etiqueta” um pouco menos problemática.

Motivado pelo trabalho de Gasque de Souza (2021) e, principalmente, pelo trabalho de Kriech (2022), um dos objetivos de Rezende dos Reis (2023) consistiu em realizar uma análise acústica qualitativa preliminar em um conjunto de dados do Português do Brasil e do Português da Europa (PE), na tentativa de verificar, mediante a observação da modulação da Frequência Fundamental (Fo) dos DPs redobrados, indícios prosódicos que pudessem sinalizar, em especial na variedade brasileira (carioca), a existência dos 2 tipos (sintáticos) de redobro do sujeito, anteriormente mencionados. Na aplicação do critério acústico apontado por Kriech (2022) para a verificação da posição do DP, foi possível identificar que o PB demonstra evidentes casos de DE, nos quais o movimento melódico dos DPs à esquerda forma tanto curvas de Fo descendentes quanto ascendentes (cf. Figura 3); tal resultado vai ao encontro de trabalhos anteriores acerca do comportamento acústico de um DP tópico (cf. Callou *et alii*, 2003 [1993] Orsini, 2003; Silva, 2018; entre outros).

Figura 3 – Modulação de Fo, transcrição fonética e transcrição ortográfica da sentença do PB “[minha mãe]... ela, queria que...”

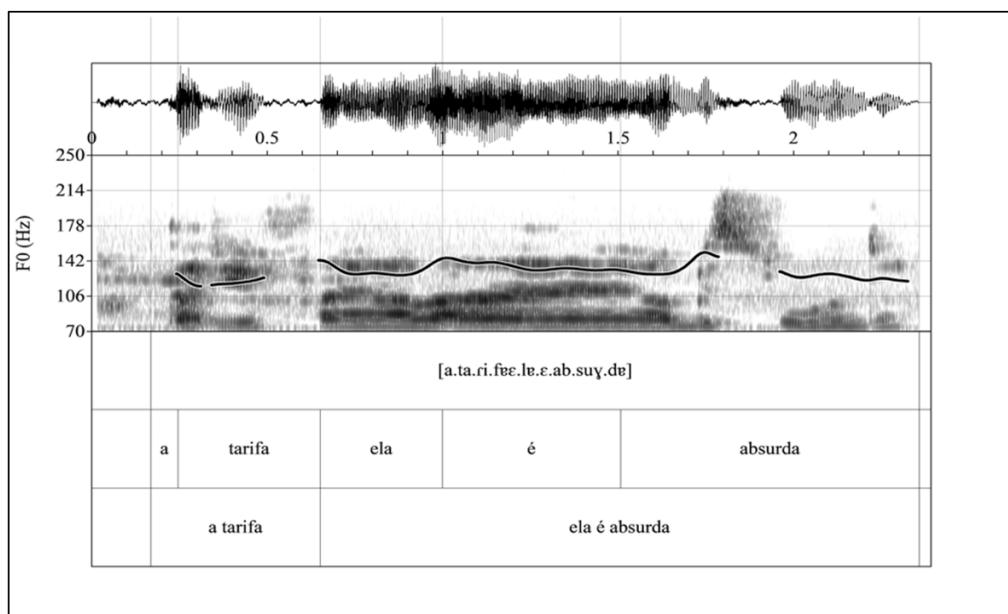


Fonte: adaptado de Rezende dos Reis (2023, p. 116).

entanto, pouco elucidativa do ponto de vista prosódico e entoacional, o que reforçaria a necessidade de que fosse realizado igualmente um tratamento acústico em seus dados.

Ao se deparar, no entanto, com a modulação de F_0 presente na Figura 4, Rezende dos Reis (2023) chamou a atenção para o movimento melódico alinhado ao DP inicial “a tarifa”, que revela uma curva com menor variação de F_0 , se comparada àquela identificada na Figura 3. Com efeito, verifica-se uma maior integração prosódica entre o DP linearmente à esquerda e a sentença que o segue, um comportamento que se distancia do que se tem definido como uma curva “prototípica” de um DP tópico (cf. Orsini, 2003).

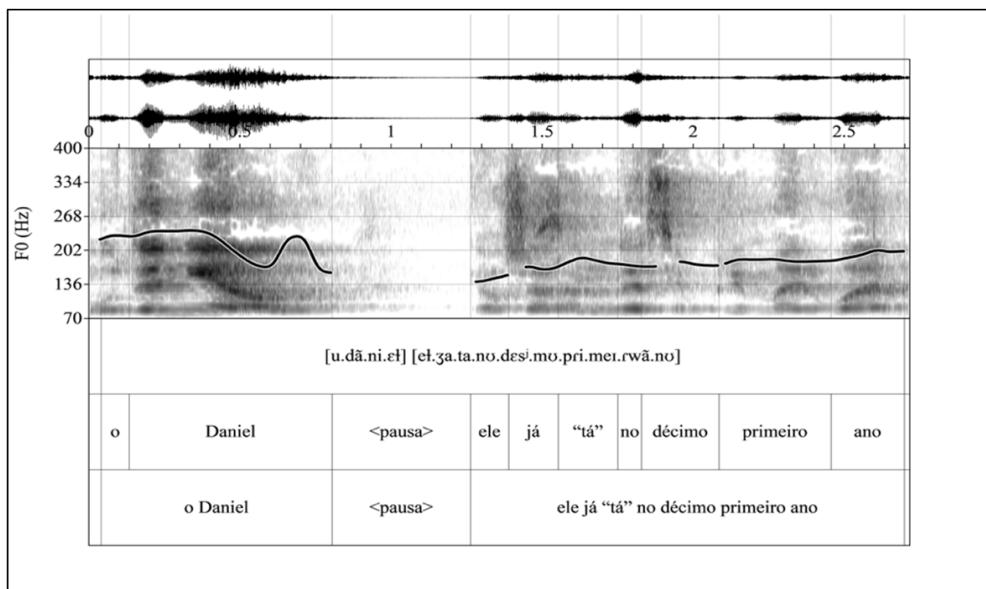
Figura 4 – Modulação de F_0 , transcrição fonética e transcrição ortográfica da sentença do PB “[a tarifa], ela, é absurda”



Fonte: adaptado de Rezende dos Reis (2023, p. 118).

No que se refere ao PE, as (poucas) ocorrências de redobro são constituídas de DPs em posição inicial cujas curvas melódicas se enquadram no comportamento prototípico de um constituinte tópico (cf. Barros, 2014; Orsini, 2003; Yano; Fernandes, 2020), conforme ilustrado na Figura 5. Rezende dos Reis (2023) conclui, então, que esses DPs são fraseados em sintagmas entoacionais (IPs) independentes da sentença-comentário, com fronteira prosódica marcada, acompanhada de pausa, o que valida a interpretação de que configuram casos emblemáticos de DEs.

Figura 5 – Modulação de F0, transcrição fonética e transcrição ortográfica da sentença do PE “[o Daniel],... ele, já ‘tá’ no décimo primeiro ano”



Fonte: adaptado de Rezende dos Reis (2023, p. 117).

Com a finalidade de refinar a análise piloto de Rezende dos Reis (2023), este trabalho tem como objetivo descrever marcas prosódicas das construções de redobro do sujeito no PB. Amparado pela hierarquia Prosódica (cf. Nespor; Vogel, 2007 [1986]) e pela Fonologia Entoacional (cf. Ladd, 2008 [1996]), argumento que a ocorrência de fronteira prosódica entre o DP inicial e o pronome resumptivo configura um traço fonológico que permite tanto sinalizar quanto diferenciar realizações prosódicas relacionadas ao redobro do sujeito. Quanto à hipótese, advogo em favor da existência de 2 estruturas prosódicas de redobro no PB, que podem ser reflexos de processos derivacionais (sintáticos) distintos (cf. Krieck, 2022; Rezende dos Reis; Duarte, 2024). Em linhas gerais, este trabalho assume que, quando o DP inicial estiver deslocado (constituindo, portanto, um verdadeiro tópico), ele será mapeado em um IP independente da sentença-comentário, conforme já apontado em outros trabalhos (cf. Silva, 2018; Yano; Fernandes, 2020); no entanto, somados a essas ocorrências, podem ainda ser observados casos em que o DP inicial é fraseado, junto da sentença que o segue, em um único IP⁶.

A amostra provém do *Corpus Concordância*, que compõe o Projeto COMPARAPORT, de onde foram extraídos dados de indivíduos cariocas. Com a finalidade de expandir a amostra, foram coletadas, da plataforma *Youtube* e do canal midiático *Globo News*, mais 30 ocorrências de redobro do sujeito para o PB, todas igualmente produzidas por falantes cariocas. Para o tratamento prosódico, foi utilizado o programa de análise acústica PRAAT

⁶ Os resultados de Silva (2018), que analisou o redobro do sujeito (chamado pela autora de DE) no gênero “debate”, evidenciam que o DP tópico e a sentença-comentário são fraseados em IPs independentes, um formado apenas pelo DP tópico e um formado pela sentença-comentário. A autora, entretanto, chama a atenção para um dado de redobro cujo comportamento destoa dos demais; o DP inicial que compõe a referida ocorrência não exibia pistas acústicas indicativas de fronteira prosódica, o que, na interpretação deste trabalho, indica que o DP e a sentença comentário se encontram fraseados em um único IP.

(Boersma; Weenink, 2023); o tratamento estatístico, por sua vez, foi realizado com o auxílio do software MINITAB versão 21.1.

Este artigo se organiza da seguinte forma: na seção 2, apresento, brevemente, os pressupostos teóricos da Fonologia Prosódica (cf. Nespor; Vogel, 2007 [1986]) e da Fonologia Entoacional (cf. Ladd, 2008); na seção 3, apresento a Metodologia, em que são explicitados a amostra analisada, o tratamento que lhe foi dado, os objetivos e a hipótese; na seção 4, reporto a análise dos dados com o redobro do sujeito, que compreende, em 4.1, a problematização dos critérios acústicos utilizados por Gasque de Souza (2021) e Krieck (2022), com o intuito de demonstrar a eficiência do novo critério aqui proposto, bem como, em 4.2, o tratamento estatístico dos dados; finalmente, teço algumas considerações sobre os resultados obtidos neste trabalho.

2 Quadro Teórico

2.1 Fonologia Prosódica

Para Nespor e Vogel (2007 [1986]), a Fonologia Prosódica é um modelo formal (de orientação gerativista) não linear para estruturas prosódicas, cuja representação abstrata se apresenta organizada hierarquicamente. Conforme salienta Tenani (2017), as autoras entendem que as relações sintáticas fornecem válidas informações na constituição e (no reconhecimento) das estruturas prosódicas (no que se refere aos algoritmos de formação desses domínios prosódicos); um dos pressupostos fundamentais dessa modelagem, no entanto, prevê que o isomorfismo entre sintaxe e prosódia não é obrigatório⁷ (cf. Nespor; Vogel, 2007 [1986], p. 2). Acrescenta-se que, nesse modelo, os domínios prosódicos se organizam da camada mais baixa à mais alta, da seguinte forma: (a) sílaba (Syl), (b) pé métrico (F), (c) palavra prosódica (PW), (d) grupo clítico (CG), (e) sintagma fonológico (PhP), (f) sintagma entoacional (IP) e (g) enunciado fonológico (U). A projeção de uma unidade de um domínio mais alto prevê, desse modo, a concatenação de unidades de um domínio mais baixo.

Ao estudo entoacional das construções aqui investigadas, interessa, sobretudo, o domínio do IP, em virtude de ser esse o domínio prosódico que viabiliza a averiguação de propriedades fonético-fonológicas recorrentemente atribuídas a um DP tópico (cf. Silva, 2018; Yano; Fernandes, 2020). De acordo com Nespor e Vogel (2007 [1986], p. 188), o IP é definido como o domínio de um contorno entoacional, em cuja fronteira é possível a inserção de uma pausa melódica. Na concepção deste trabalho, o nível do IP corresponde ainda à camada que permite identificar um comportamento prosódico (e entoacional) das construções de redobro do sujeito que se distancia do “prototípico” observado por Orsini (2003) e por Silva (2018).

⁷ Há pelo menos 25 anos, a literatura de base gerativista dispõe da já mencionada abordagem cartográfica (cf. Cinque; Rizzi, 2010), que, apesar de “sintaticocêntrica” (cf. nota 4), possibilita “estreitar as relações” entre sintaxe, prosódia e discurso, pelo menos no que diz respeito a fenômenos relacionados à periferia esquerda da sentença. Um ótimo exemplo é o trabalho de Frascarelli e Hinterhözl (2007), no qual os autores postulam uma tipologia tripartite para a categoria “tópico”; nela, cada subtipo de tópico ocupa uma posição sintática distinta na borda esquerda da sentença, bem como apresenta propriedades entoacionais e discursivas próprias.

2.2 Fonologia Entoacional

Para este artigo, são assumidos igualmente os pressupostos da Fonologia Entoacional, em seu modelo Métrico-Autossegmental - MA - (cf. Pierrehumbert, 1980; Ladd, 2008 [1996]; entre outros). Com base nesse modelo, é possível entender a entoação como um arranjo fonológico associado “ao uso de características fonéticas suprasegmentais para transmitir significados ‘pós-lexicais’ ou significados pragmáticos no nível da sentença de uma forma linguisticamente estruturada” (cf. Ladd, 2008, p. 4)⁸. Assume-se que a entoação corresponde a uma representação abstrata própria - ou seja, independente de outros fenômenos fonológicos (cf. Serra, 2009). Quanto à sua composição, o referido expediente fonológico é formado de eventos tonais discretos, responsáveis por sinalizar pontos linguisticamente importantes na estrutura tonal.

De acordo com Pierrehumbert (1980), os eventos tonais que se alinham às sílabas prominentes são chamados de “acentos tonais” (*pitch accents*), enquanto aqueles que indicam a fronteira de um constituinte prosódico recebem o rótulo de tons de fronteira (*boundary tones*). Para a descrição desses eventos, costuma-se adotar uma notação (abstrata) composta de 2 tons, o tom alto (H) e o tom baixo (L), que dão origem a padrões entoacionais mono ou bitonais. Caso tenha de sinalizar uma sílaba tônica, o tom é seguido de um asterisco sobrescrito (L*; H*); caso demarque a fronteira de um sintagma entoacional, o tom é seguido do símbolo “%”. Tal aparato descritivo, composto de um inventário de tons e de diacríticos que descrevem os eventos tonais, anos mais tarde, ficou conhecido como notação *ToBI*. Em síntese, o modelo MA, brevemente descrito, tem como objetivo analisar e descrever a estrutura entoacional das línguas, mediante o mapeamento dos contrastes entre os eventos tonais, buscando, por intermédio de uma notação mínima, chegar a possíveis padrões entoacionais universais.

3 Metodologia

A análise se valeu de 55 dados extraídos da Amostra Concordância, que compõe o Projeto COMPARAPORT: Estudo Comparado dos Padrões de Concordância em Variedades Africanas, Brasileiras e Europeias do Português (Projeto de cooperação entre a UFRJ e o Centro de Linguística da Universidade de Lisboa, disponível em www.corporaport.ufrj.br). A referida amostra consiste de entrevistas gravadas entre 2008 e 2010, com falantes de Lisboa (Oeiras e Cacém) e Rio de Janeiro (Copacabana e Nova Iguaçu), estratificados segundo a faixa etária (18-35 / 36-55 / 56-75), o nível de escolaridade (Ensino Fundamental, Médio e Superior) e o gênero. Em adição a esse cômputo, com a finalidade de aumentar a amostra, foram coletadas, da plataforma *Youtube* e do canal midiático *Globo News*, mais 30 ocorrências de redobro do sujeito para o PB, todas igualmente produzidas por falantes cariocas.

Como recorte metodológico, a análise focalizou estruturas com o redobro de DPs lexicais em sentenças declarativas; foram consideradas apenas as ocorrências em que DPs iniciais se encontram vinculados a pronomes nominativos de 3^a pessoa na função de sujeito sintático,

⁸ Do original: “[...] to the use of suprasegmental phonetic features to convey ‘postlexical’ or sentence-level pragmatic meanings in a linguistically structured way.”

sem que haja rupturas ocasionadas por material interveniente, como nos exemplos em (2a-b). Tal decisão metodológica se justifica dada a tentativa de evitar que o contorno entoacional do constituinte interpolado interfira em uma averiguação mais precisa da modulação de Fo do DP linearmente à esquerda.

- (2) a. [a doméstica], *ela*, trabalha trabalha trabalha.
b. [a imprensa],...⁹ *ela*, vive disso.

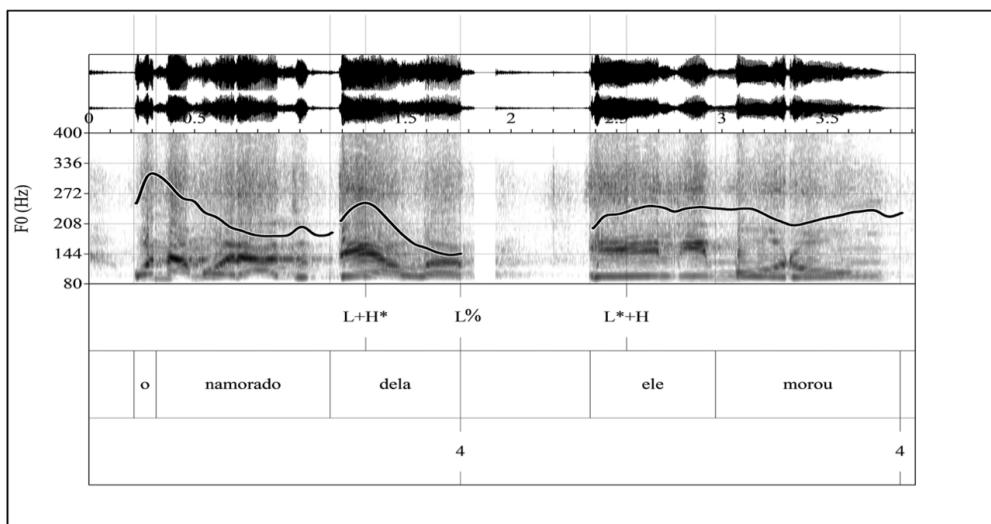
Partindo das considerações prévias obtidas por Rezende dos Reis (2023), este trabalho tem como objetivo geral descrever (algumas) marcas prosódicas das construções de redobro do sujeito. Amparado pela hierarquia Prosódica, nos moldes de Nespor e Vogel (2007 [1986]), e pelo modelo Métrico-Autossegmental (Ladd, 2008 [1996]), defendo que a presença ou a ausência de fronteira prosódica entre o DP inicial e o pronome resumptivo que o retoma consiste em um traço fonológico que permite tanto sinalizar quanto diferenciar realizações prosódicas referentes ao redobro do sujeito. Para tanto, considerei observar pistas acústicas de ordem duracional e melódica, que indicassem a presença dessa fronteira, com substantivo destaque para o controle da ocorrência da pausa (cf. Tenani, 2002; Serra, 2009). Quanto à hipótese, argumento que haja 2 realizações prosódicas associadas ao fenômeno do redobro do sujeito no PB, possíveis reflexos de algoritmos derivacionais (sintáticos) distintos, se nos apoiarmos em trabalhos como o de Kriech (2022). Em linhas gerais, entendo que, quando o DP inicial estiver deslocado, seu fraseamento prosódico ocorrerá em um IP independente da sentença-comentário; no entanto, haverá ainda ocorrências em que o DP em posição inicial e a sentença subsequente serão mapeados em um único IP. Na intenção de medir se o fraseamento do DP à esquerda em um IP independente é condicionado pelo seu peso fonológico, foi controlado igualmente o número de PWs¹⁰ que constitui esse DP.

Para o exame das propriedades acústicas, foi utilizado o software PRAAT (Boersma; Weenink, 2023). A notação dos dados apresentados se valeu do sistema P_TOBI (cf. Frota *et al.*, 2015), que, neste artigo, exibirá 3 tiers, alinhados a pontos específicos do contorno de Fo, conforme se verifica na Figura 6: (a) o evento tonal e o tom de fronteira (ou seja, o contorno nuclear) do DP em posição inicial (em caso de presença de fronteira) e o evento tonal do pronome resumptivo; (b) a transcrição ortográfica de cada palavra dicionarizada da sentença; (c) a anotação de fronteiras prosódicas (especificamente do IP).

⁹ Neste artigo, a presença de pausa será sinalizada por meio do recurso gráfico “reticências” (...).

¹⁰ Em linhas gerais, o domínio de PW é aquele em que é mapeado apenas um acento primário (cf. Nespor; Vogel, 2007 [1986]).

Figura 6 – Modulação de F0, notação entoacional, transcrição ortográfica e fraseamento da sentença do PB “[o namorado dela],... *ele*, morou...”



Fonte: Elaboração própria.

O tratamento estatístico foi realizado com o auxílio do software MINITAB versão 21.1, por meio do qual se processou uma modelagem de regressão logística. Nesse sentido, a referida modelagem permite estimar os efeitos das variáveis preditoras (ou independentes) na tomada de decisão por uma das “faces” da variável resposta (ou dependente), que é obrigatoriamente de natureza binária. Dado o importante papel da fronteira prosódica neste trabalho, tal expediente fonológico foi tomado como variável resposta; as variáveis preditoras¹¹ medidas foram:

- (a) “ocorrência de pausa” - sim ou não;
- (b) “número de PWs do DP inicial” - 1, 2 e 3 (ou mais)¹²;

Com a modelagem, esta investigação buscou, então, verificar se as variáveis independentes elencadas figuram como pistas prosódicas significantes (do ponto de vista estatístico), que permitiriam evidenciar (e até mesmo prever) a presença de fronteira prosódica em construções de redobro do sujeito.

Em resumo, este artigo tem 2 objetivos específicos: (a) argumentar em favor de um critério de reconhecimento do tipo do redobro pautado na averiguação da fronteira prosódica; (b) fornecer um tratamento estatístico para medir a atuação de parâmetros acústicos na marcação de fronteira.

¹¹ Um dos pressupostos da regressão logística é o de que as variáveis preditoras devem ser independentes entre si, i.e., uma variável preditora não pode interferir no comportamento de outra preditora. Segundo Montgomery *et alii* (2006), um meio de verificar uma possível multicolinearidade entre tais variáveis é mediante a observação dos coeficientes dos “fatores de inflação da variância” (VIF). Na modelagem realizada, os coeficientes VIF obtidos para as referidas variáveis revelaram que uma possível interferência entre elas é estatisticamente irrelevante.

¹² Foi controlada ainda mais uma variável preditora, o “número de sílabas da(s) PW(s) à esquerda”, com o objetivo de verificar, tal como “o número de PWs”, o efeito do peso fonológico sobre a ocorrência de fronteira. Por limitação de espaço, os resultados para essa variável não serão reportados.

4 Análise dos dados

Esta seção se encontra dividida em duas partes. Na subseção 4.1, pretendo problematizar, pautado nos resultados aqui obtidos, os critérios acústicos usados por Gasque de Souza (2021) e Krieck (2022), que, até certo ponto, não se mostram muito precisos na diagnose de diferentes estruturas prosódicas para o redobro do sujeito no PB; busco mostrar, com o suporte de um modelo fonológico, que o reconhecimento de uma fronteira prosódica parece configurar um critério mais eficaz na distinção do comportamento prosódico das referidas construções. Na subseção 4.2, reporto os resultados para o tratamento estatístico dos dados, que teve como finalidade verificar o modo pelo qual a ocorrência de fronteira prosódica se “articula” a (ou pode decorrer de) outras propriedades fonético-fonológicas, a ocorrência de pausa e o número de PWs do DP em posição inicial.

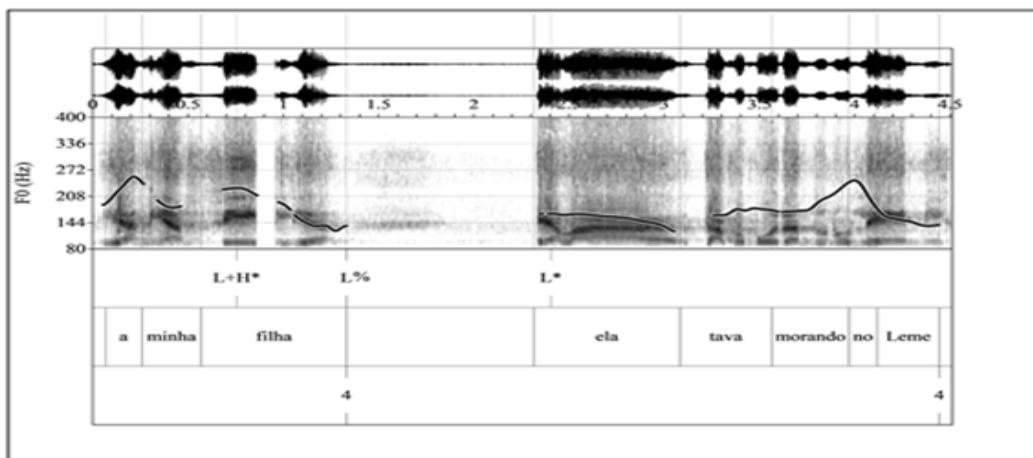
4.1 Por um critério mais abstrato: a ocorrência de fronteira melódica

Conforme apresentado na seção 1, há trabalhos que assumem que as construções abrigadas sob o termo “redobro do sujeito” não se comportam de modo uniforme. Neste sentido, a heterogeneidade observada nas referidas construções pode sugerir que não estamos diante de estruturas que constituem um mesmo conjunto. No PB, haveria, portanto, estruturas sintáticas de redobro do sujeito com DPs iniciais efetivamente deslocados, bem como aquelas cujos DPs ocupariam uma posição hierarquicamente mais baixa, se comparada à de um DP tópico. Para validar tal argumentação, alguns estudos tendem a recorrer a aspectos prosódicos, associados sobretudo ao DP em posição inicial, na defesa de que cada tipo de redobro apresentaria algoritmos derivacionais (sintáticos) próprios, cada um dos quais a apresentar traços (não somente) acústicos particulares (cf. Krieck, 2022). Gasque de Souza (2021), por exemplo, entende que a ocorrência do parâmetro acústico “pausa” entre o DP inicial e o pronome resumptivo, combinado com o estatuto informacional desse DP, atua como um traço que diferenciaria as “DEs verdadeiras” das “falsas”. Krieck (2022), por sua vez, concebe que o contorno entoacional do DP inicial, acompanhado ou não de pausa, corresponde a uma marca que indica em que posição, na arquitetura da sentença, tal DP se encontra. No entanto, até então, nenhum trabalho que assume a posição ora apresentada fundamentou seu empreendimento em um modelo fonológico (ou em abordagens que se inserem nesse modelo).

Amparando-me na hierarquia prosódica de Nespor e Vogel (2007 [1986]) e na Fonologia Entoacional, em seu modelo Métrico-Autossegmental, argumento, nesta subseção, que um critério mais preciso para identificar diferentes estruturas prosódicas associadas ao redobro no PB consiste em considerar uma propriedade mais abstrata, a ocorrência de fronteira prosódica, alinhada, se presente, a um tom de fronteira alto ou baixo. Dito de outra forma, defendo que um diagnóstico mais acurado decorre da apuração do fraseamento prosódico do DP em posição inicial nas construções de redobro. Em trabalhos que focalizam aspectos prosódicos e entoacionais da categoria “tópico”, recorrentemente se destaca que um DP deslocado (ou seja, um DP tópico) é mapeado em um IP independente da sentença comentário (cf. Silva, 2018; Yano; Fernandes, 2020), cuja fronteira se reconhece por meio de determinadas pistas acústicas, como a pausa (cf. Nespor; Vogel, 2007 [1986]; Tenani, 2002;

Serra, 2009, entre outros). Entre os dados examinados, foram encontradas ocorrências de redobros que confirmam a literatura especializada (de que um DP tópico é fraseado em um IP independente). Vejamos a Figura 7¹³:

Figura 7 – Modulação de Fo, notação entoacional, transcrição ortográfica e fraseamento da sentença do PB “[a minha filha],... *ela*, tava morando no Leme...”



Fonte: Elaboração própria.

Na Figura 7, observa-se uma curva entoacional descendente, referente ao DP “a minha filha”, cuja última PW (“filha”) se alinha a um contorno nuclear L+H* L%, seguido de pausa; o pronome resumptivo, que retoma o DP tópico, se alinha a um acento tonal baixo (L*). A modulação de Fo observada (do DP linearmente à esquerda) possivelmente se enquadra no que Kriech (2022) denomina de “contorno de tópico”. Seria, no entanto, tal comporta-

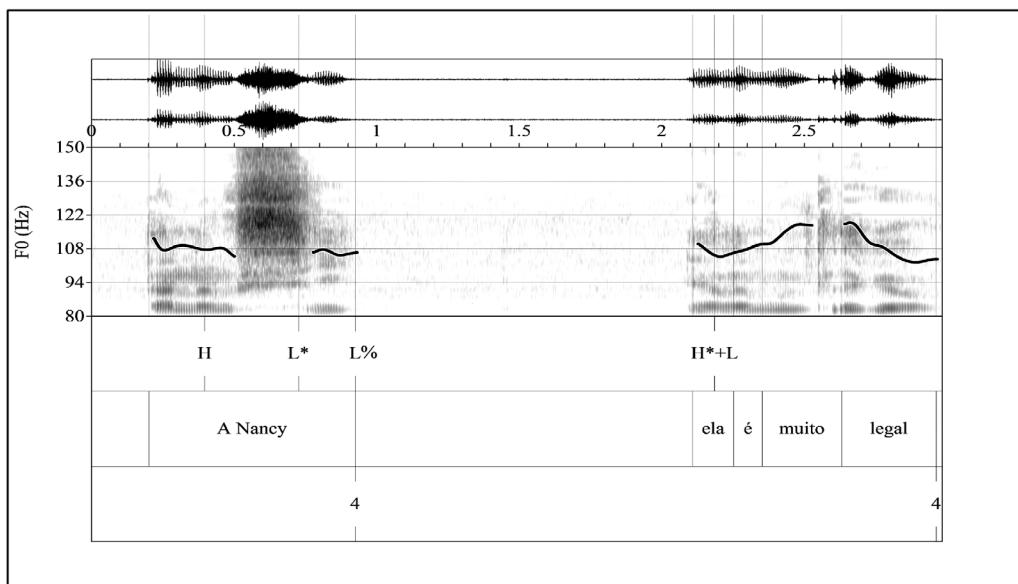
¹³ Ainda que não seja o objetivo primeiro deste trabalho rastrear uma correlação entre contornos entoacionais e funções discursivas, abro parênteses para tecer algumas considerações sobre o exemplo apresentado na Figura 7, que, quando contextualizado (i), pode fornecer interessantes informações sobre determinadas características do redobro do sujeito presentes no PB.

(i) “L: Meu filho morou até o ano passado comigo, mas, ele trabalha na Barra, na Rio Esporte Center. Ele é *personal trainer*. Então, ele resolveu mudar para o lado de lá. [...] Aí, ele agora tá morando no Recreio. Tá, tá muito mais perto. **[a minha filha]** *ela*, **mora no Leme.**”

Veja que, em (i), é articulada uma relação de contraste por oposição entre a informação codificada pelo DP inicial na construção com o redobro e a informação articulada anteriormente (cf. Rosa-Silva, 2019); a participante da entrevista contrasta o fato de o seu filho morar no Recreio e a sua filha, alvo do redobro, no Leme. E esse contraste, na construção com o redobro, é marcado por uma modulação de Fo do DP inicial bastante característica, demarcada por uma fronteira prosódica. Ao investigar, na amostra, outros dados de mesma natureza, parece haver uma tendência de a relação de contraste, que, em (i), se estabelece entre os filhos da participante, estar associada a uma entonação similar à da Figura 7, o que parece não ser recorrente em enunciados que não codifiquem tal função. Soma-se a isso a assunção de que, segundo Frascarelli e Hinterhölzl (2007; cf. nota 7), DPs com o estatuto de tópico contrastivo, que é o caso do DP em exame, congelam em uma posição sintática específica na borda esquerda da sentença. A partir de tal mapeamento, temos em mãos alguns indícios de uma relação entre “discurso” (contraste), prosódia (modulação específica com fronteira prosódica demarcada) e sintaxe (DP em uma posição A’), o que fortalece a tese de Kriech (2022), segundo a qual uma construção de redobro com um DP efetivamente deslocado apresentaria um percurso derivacional próprio, com reflexos nos níveis de interface. Naturalmente, tal mapeamento, que investigará não somente a relação de contraste, terá mais destaque em trabalhos futuros.

mento prosódico *per se* uma condição indispensável para a validação de um DP como tópico? Veja-se a Figura em 8:

Figura 8 – Modulação de Fo, notação entoacional, transcrição ortográfica e fraseamento da sentença do PB “[a Nancy],... ela, é muito legal...”



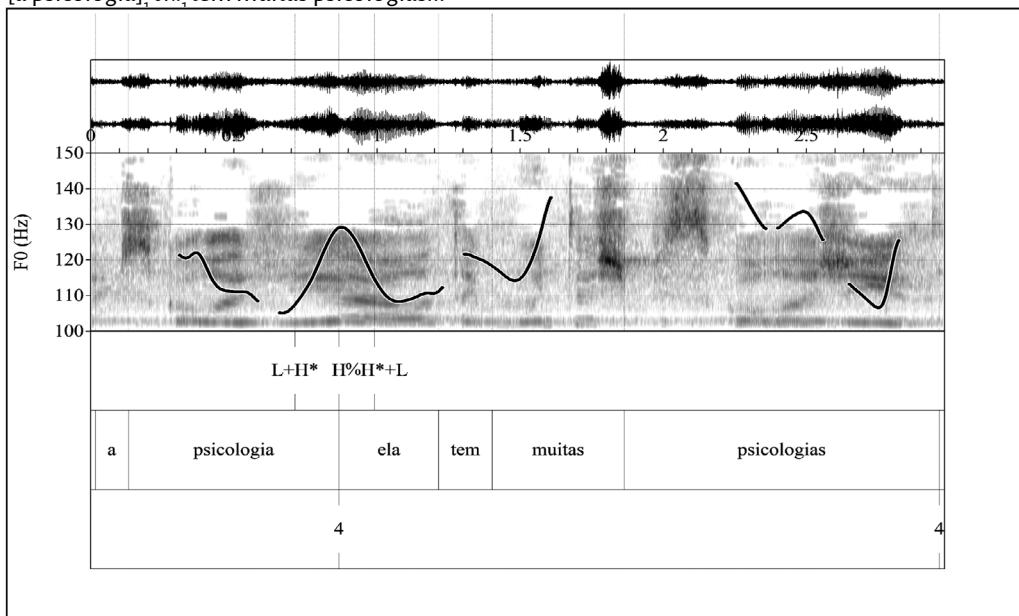
Fonte: Elaboração própria.

Ao analisar a Figura 8, visualiza-se um DP inicial, cuja linha de Fo permanece “contínua”, sem apresentar modulações marcadas que a literatura comumente associa a um DP tópico (cf. Orsini, 2003). Mesmo na ausência de uma “curva prototípica de tópico”, o referido DP inicial, inegavelmente um tópico marcado, apresenta uma fronteira baixa, que se segue de uma pausa silenciosa; é fraseado, portanto, em um IP independente¹⁴. Nesse cenário, o que parece garantir a sua interpretação (prosódica) como um DP tópico é um traço fonológico, a presença de fronteira prosódica (não uma eventual modulação de Fo), sinalizada, na ocorrência em exame, pela presença da pausa.

Nos dados analisados, ainda foram encontrados casos em que, na ausência de pausa, a fronteira prosódica era sinalizada por intermédio de uma outra pista acústica, o que parece validar o DP inicial como um verdadeiro tópico; corresponde, portanto, a um possível contraexemplo à proposta de Gasque de Souza (2021), tendo em vista que estaríamos diante da ocorrência de uma DE legítima. A título de exemplificação, vejamos a Figura 9:

¹⁴ Ao averiguar algumas ocorrências que não foram formalmente contempladas neste artigo (aqueles em que há material interveniente ao DP tópico e o pronome que a ele se vincula), pude identificar casos em que o DP tópico se comporta prosodicamente de modo similar ao da Figura em questão - ou seja, apresenta uma modulação de Fo menos marcada. Diante desse cenário, surge um questionamento: na presença de uma outra pista que sinalize a existência de uma fronteira melódica (como, por exemplo, a ocorrência de pausa, ou a presença de um elemento interventor), o contorno de Fo “típico de tópico” se converteria em uma “estratégia” opcional desse DP deslocado?

Figura 9 – Modulação de Fo, notação entoacional, transcrição ortográfica e fraseamento da sentença do PB “[a psicologia], *ela*, tem muitas psicologias...”

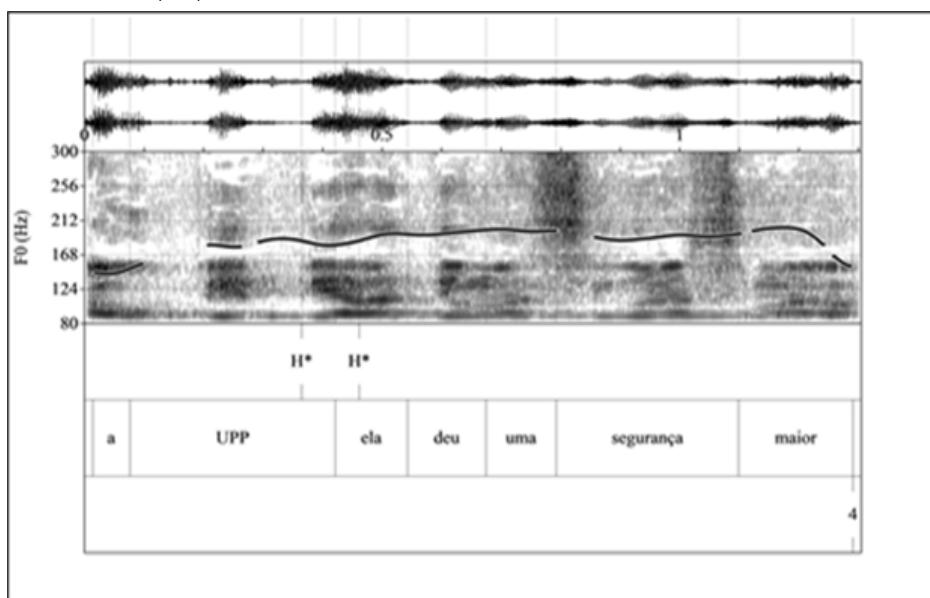


Fonte: Elaboração própria.

Ao examinar a Figura em destaque, visualiza-se que o contorno nuclear do DP tópico forma um padrão $L+H^*$ $H\%$; a gama da variação de Fo - uma queda brusca de Fo que ocorre após o contorno do DP “a psicologia”; ou seja, a curva descendente alinhada ao evento bitonal H^*+L - parece indicar a existência de uma fronteira prosódica.

Para além dos casos de redobro em que se verifica um constituinte à esquerda fraseado em um IP independente, foram encontrados dados em que o DP inicial apresenta uma maior integração prosódica com a sentença que se segue, casos específicos (e intrigantes) já anteriormente pontuados por Gasque de Souza (2021) e Rezende dos Reis (2023). Na Figura 10, dispõe-se uma dessas ocorrências.

Figura 10 – Modulação de Fo, notação entoacional, transcrição ortográfica e fraseamento da sentença do PB “[a UPP], *ela*, deu uma segurança maior”



Fonte: Elaboração própria.

Embora ocorra uma leve ascensão da curva de Fo, alinhada ao determinante “a” e à primeira sílaba pretônica do acrônimo “UPP”, o restante do contorno se mantém “contínuo”, pouco marcado, sobretudo se comparado à modulação do DP inicial presente na Figura 7. Verifica-se, então, que tanto o acento tonal do DP inicial quanto o do pronome-cópia são altos (H*). O que mais chama a atenção, contudo, é justamente a ausência de quaisquer pistas indicativas de uma fronteira prosódica, um comportamento que, na interpretação deste trabalho, caracteriza esse tipo inovador de redobro do PB.

Diante do que foi exposto, considero que verificar a ocorrência de fronteira prosódica no DP em posição inicial (em vez de se limitar à verificação da ocorrência de pausa e do contorno de Fo) configura um critério mais eficiente no reconhecimento e na caracterização das realizações prosódicas das construções de redobro do sujeito. Em linhas gerais, o critério de Krieck (2022) não dá conta de possíveis DPs tópicos que não apresentam “um contorno prototípico de tópico”, como aquele ilustrado na Figura 8; o critério de Gasque de Souza (2021), por sua vez, não prevê ocorrências como a da Figura 9, casos de DEs verdadeiras que não apresentam uma ruptura na melodia entre o DP inicial e o pronome resumptivo. Um critério ancorado na verificação da fronteira prosódica entende as visões de Krieck e Gasque de Souza não como condições *sine qua non*; antes, interpreta-as como pistas para a sinalização de uma característica fonológica, concebendo que os diferentes tipos de redobro revelam estruturas entoacionais distintas. Em outras palavras, um critério pautado na verificação da ocorrência de fronteira prosódica, além de “unificar” os critérios anteriores, permite deduzir que cada uma das estruturas prosódicas de redobro apresenta uma estrutura fonológica própria.

4.2 O tratamento estatístico

O foco desta subseção recai em apresentar os resultados para o tratamento estatístico do comportamento prosódico do redobro do sujeito, tendo como parâmetro de análise a ocorrência de fronteira prosódica no DP linearmente à esquerda; pretendo mostrar, dessa forma, o modo pelo qual a referida propriedade fonológica, que se mostra ou não presente nas estruturas de redobro (cf. subseção 4.1), se “articula” a (ou pode decorrer de) outras propriedades fonético-fonológicas, a ocorrência de pausa e o número de PWs do DP em posição inicial. Na Tabela 1, encontram-se distribuídos os dados de redobro do sujeito, segundo a ocorrência de fronteira prosódica, a ocorrência de pausa e o número de PWs do DP redobrado.

Tabela 1 – Distribuição de ocorrências de redobro do sujeito (ocorrência de fronteira prosódica), segundo a ocorrência de pausa e número de PWs do DP inicial

	AUSÊNCIA DE FRONTEIRA		PRESENÇA DE FRONTEIRA	
	Sem Pausa	Com Pausa	Sem Pausa	Com Pausa
1 PW	19 / 43	--	5 / 12	5 / 30
2 PWs	12 / 43	--	3 / 12	16 / 30
≥ 3 PWs	12 / 43	--	4 / 12	9 / 30
	43 (51%)		42 (49%)	

Fonte: Elaboração própria.

Ao verificar a Tabela 1, observa-se que a distribuição das ocorrências de redobro, no que concerne à ocorrência de fronteira prosódica, é equilibrada. Do total analisado, visualiza-se que 43 sentenças não apresentam indícios acústicos responsáveis por demarcar a fronteira de IP entre o DP inicial e o pronome resumptivo, o que sinaliza que esse DP se encontra mais integrado prosodicamente à sentença que o segue (cf. Figura 10). A distribuição dos dados ainda sugere que o peso fonológico não opera necessariamente como uma restrição a tal integração, tendo em vista que 24 dos 43 casos de redobro sem fronteira prosódica são constituídos de DPs iniciais com 2 ou mais PWs.

No conjunto de dados com fronteira melódica entre o DP e o pronome-cópia, também se atesta uma distribuição pertinente. Mesmo que a presença de pausa, nos dados analisados, consista em uma pista acústica de grande protagonismo na demarcação de fronteira de IP, estando presente em 30 dos 42 dados (71%), o referido expediente não se mostrou uma condição *sine qua non* na identificação desse ponto limítrofe. Em 12 dos 42 dados, a estratégia verificada na indicação da fronteira prosódica consiste na gama da variação de Fo do DP inicial (cf. Figura 9). Uma outra informação, que chama a atenção no montante de ocorrências evidenciado, se apresenta na predominância de casos com pausa nos quais o DP inicial é

constituído de 2 ou mais PWs, 25 de 30 dados (83%), o que pode indicar que o peso fonológico teria algum efeito sobre a presença de fronteira¹⁵.

Com o intuito de estimar a relevância estatística das propriedades fonético-fonológicas aqui analisadas em relação à variável “ocorrência de fronteira prosódica”, foi realizada uma modelagem de regressão logística, que teve como valor de aplicação o fator de interesse “com fronteira”. No Quadro 1, dispõe-se a “análise da variância” (ou ANOVA)¹⁶, uma das medições realizadas em conjunto com tal modelagem.

Quadro 1 – Coeficientes obtidos na modelagem de regressão logística

Fonte	Desv (Aj.)	Média (Aj.)	Qui-Quadrado	Valor-P
Regressão	66,2	13,25	66,26	0,00
Pausa	55,4	55,47	55,47	0,00
PW	0,0	0,00	0,01	0,99
Erro	58,2	0,72		
Total	124,5			

Fonte: Elaboração própria.

Dos coeficientes fornecidos no Quadro 1, o que nos importa é o valor de probabilidade, referenciado como *valor-p*, que consiste em uma medida estatística de verificação da *hipótese nula*. Neste trabalho, a hipótese nula (de natureza estatística, ou H_0) corresponde à afirmação de que as variáveis preditoras “pausa” e “número de PWs” não teriam efeito na variável resposta “ocorrência de fronteira”. Nos manuais de estatística, tende-se a convencionar como *valor-p* mínimo a medida 0,05¹⁷ (cf. Guy; Zilles, 2007; Oliveira, 2009), abaixo da qual já é possível rejeitar, em termos estatísticos, a H_0 . Com base nisso, ao focalizar a variável preditora “número de PWs”, com *valor-p* “0,99” > 0,05, fica evidente que essa variável independente não se apresenta como estatisticamente relevante para a variável resposta “ocorrência de fronteira”¹⁸, contrariando, de certo modo, a impressão primeira obtida na observação da Tabela 1. Em contrapartida, visualiza-se que o expediente acústico “pausa”, que exibe um *valor-p* < 0,01 (o melhor cenário possível), se associa substancialmente ao fator de interesse em exame, a presença de fronteira. Para que fique mais clara a relação entre essas duas variáveis, será

¹⁵ Num primeiro momento, em que ainda não havia sido feita a modelagem de regressão logística, considerei que o peso fonológico poderia estar atuando diretamente na presença de pausa, o que fere um dos pressupostos de tal modelagem. No entanto, a partir dos coeficientes VIF (cf. nota 11), pôde ser verificada a ausência de multicolinearidade.

¹⁶ De acordo com Gujarati e Porter (2011, p. 289), a ANOVA é utilizada com a finalidade de “[...] avaliar o significado estatístico da relação entre um regressando quantitativo e regressores binários ou qualitativos”.

¹⁷ Tal valor-p deve ser compreendido da seguinte forma: em um mundo hipotético, há 5% de probabilidade de que a hipótese nula seja verdadeira. Na área da Econometria, contudo, são normalmente consideradas 3 medidas de *valor-p*, 0,1 - 0,05 - 0,01; o melhor cenário se verifica, com efeito, quando o *valor-p* é $\leq 0,01$.

¹⁸ Conforme apontado na nota 12, foi igualmente medida a variável “número de sílabas da(s) PW(s)” à esquerda. No entanto, tal como a variável “número de PWs”, ela não se mostrou relevante estatisticamente, o que pode sugerir que o peso fonológico, nos dados tratados, não se apresenta como fundamental para a variável resposta. Ressalto, contudo, que este tratamento estatístico (ainda preliminar) será, naturalmente, refeito em trabalhos futuros, quando a amostra for expandida.

demonstrado, no Quadro 2, o resultado referente à medida de efeito *odds ratio* (a razão de chances), um coeficiente de associação (entre fatores de uma variável, i.e. uma razão entre duas possibilidades), que tem como ponto de corte o valor 1¹⁹.

Quadro 2 – Coeficiente obtido para a *odds ratio*

Nível A	Nível B	Razão de Chances	IC de 95%
Pausa			
S	P	0,0270	(0,0080; 0,0916)

Fonte: Elaboração própria.

Para a interpretação do Quadro 2, a leitura que deve ser feita é a seguinte: a chance de que um dado de redobro com fronteira prosódica ocorra sem pausa (indicado pelo código “S”) é significativamente menor do que a de que ocorra acompanhada de pausa (indicado pelo código “P”), algo da ordem de 37 vezes menos provável²⁰; esse resultado ratifica, sustentado em uma análise estatística, os trabalhos na literatura que apontam ser a pausa a estratégia mais comum na marcação de fronteira de IP (cf. Serra, 2009).

5 Considerações (nada) finais

Este artigo teve como objetivo geral mapear a(s) realização(ões) prosódica(s) da(s) construção(ões) de redobro do sujeito. Como um dos objetivos específicos, busquei argumentar que os critérios usados por Gasque de Souza (2021) e Krieck (2022), pautados na verificação da ocorrência de pausa e na observação do contorno entoacional, respectivamente, se mostram, até certo ponto, não muito precisos na diagnose de diferentes estruturas prosódicas para o redobro do sujeito. Isso se dá, na verdade, uma vez que tais parâmetros atuam “somente” como pistas (e não como condições *sine qua non*), que sinalizam uma característica mais abstrata (fonológica), a ocorrência de fronteira; esta, se presente, pode ser demarcada por um tom baixo ou alto. Um critério pautado no reconhecimento da fronteira prosódica é, portanto, mais eficaz, em virtude de que, além de “unificar” os critérios anteriores, tendo uma maior cobertura empírica, permite deduzir que cada uma das estruturas prosódicas de redobro apresenta uma estrutura fonológica própria (cf. seção 4.1). A partir do reconhecimento da fronteira prosódica, puderam ser identificadas 2 realizações prosódicas / entoacionais distintas de redobro(s) do sujeito: (a) aquela com fronteira prosódica, em que o DP inicial é clara-

¹⁹ Há 3 cenários possíveis na interpretação do valor da *odds ratio* atribuído a um determinado fator: (a) se a *odds ratio* de um fator X de uma variável preditora (em relação a um fator Y) for = 1, pode-se dizer que não há diferença na chance de ocorrência de um fator X em comparação com um fator Y; (b) se a *odds ratio* de um fator X de uma variável preditora for > 1, entende-se que X apresenta mais chances de ocorrer do que Y; (c) se a *odds ratio* de um fator X de uma variável preditora for < 1, entende-se que X apresenta menos chances de ocorrer do que Y.

²⁰ Em linhas gerais, o cálculo utilizado para se chegar ao referido índice consiste em dividir o valor de corte 1 pelo valor da *odds ratio*; verifica-se, então, que $1 / 0,027 = 37,037$.

mente fraseado em um IP independente da sentença-comentário; (b) aquela sem fronteira prosódica, em que o DP inicial e a sentença-comentário são mapeados em um único IP, apresentando uma maior integração prosódica.

Com a realização do tratamento estatístico, o segundo objetivo específico, ficou evidente que, nos dados analisados, a pausa tem um papel de substantiva importância na sinalização da fronteira prosódica, o que confirma trabalhos anteriores (cf. Serra, 2009). Por meio da medida de associação *odds ratio*, verificou-se que o redobro com fronteira, ou seja, casos em que o DP inicial é fraseado em um IP independente, acompanhado de pausa, tem bem mais “chances” de ocorrer do que o redobro com fronteira indicada por outra pista prosódica. Em contrapartida, “o número de PWs” que constituem o DP inicial não tem efeito no fator de interesse “presença de fronteira”, mostrando-se irrelevante do ponto de vista estatístico. Em outras palavras, o peso fonológico, nas construções de redobro aqui analisadas, parece não condicionar a ocorrência de fronteira prosódica. Esses resultados, portanto, podem sinalizar que, tal como apontado por Quarezmin (2019), pautada na sintaxe cartográfica do “sujeito”, não estamos diante de um único conjunto de estruturas de redobro do sujeito; antes, estão, sob essa etiqueta, diferentes “redobros”, cada um dos quais a apresentar um processo derivacional próprio (com diferentes impactos nos sistemas de interface (cf. nota 13)), o que parece confirmar a minha hipótese.

A investigação aqui apresentada consiste apenas em um primeiro passo, que espero que adiante ganhe novos contornos. Para o futuro, pretendo expandir a amostra com a qual trabalho, para que seja possível chegar a conclusões mais abrangentes; viso ainda a controlar, mais rigorosamente, outras pistas acústicas, a fim de descrever de modo mais detalhado as características acústicas e entoacionais referentes a cada realização prosódica do redobro do sujeito - nesse sentido, viso a tomar os eventos tonais, em um futuro trabalho, como uma variável a ser controlada. Finalmente, manifesto um desejo de realizar um (ou alguns) experimento(s) com as construções de redobro, em que cogito controlar fatores de diferentes ordens, com o intuito de trazer novos ingredientes que evidenciem (ou pelo menos sugiram) que os diferentes tipos (prosódicos) de redobro configuram efetivamente estruturas (sintáticas) diferentes.

Agradecimentos

Agradeço aos pareceristas *ad hoc* pelas valiosas contribuições, que inegavelmente tornaram certos pontos do texto mais claros. Agradeço igualmente ao CNPq pelo apoio financeiro, que tem permitido me dedicar, em tempo integral, às pesquisas que venho desenvolvendo. Todas as falhas remanescentes são de minha inteira responsabilidade.

Referências

- BARROS, N. *Fraseamento prosódico em Português: uma análise entoacional de enunciados com parentéticas e tópicos em duas variedades do Português Europeu*. 2014. 97 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade de Lisboa, 2014.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer [computer program]*. Versão 6.3.14. Amsterdam: Universiteit van Amsterdam; 2023. Disponível em: www.praat.org. Acesso em: 25 nov. 2023.
- BRITO, A. M.; DUARTE, I.; MATOS, G. Estrutura da frase simples e tipos de frases. In: MATEUS, M. H. M. et alii (orgs.). *Gramática da língua portuguesa*. Lisboa: Editorial Caminho, 2003. p. 433-506.
- CALLOU, D. et alii. Topicalização e deslocamento à esquerda: sintaxe e prosódia. In: CASTILHO, A. (org.). *Gramática do português falado*. Vol. III: As abordagens. 3^a ed. Campinas: Editora da Unicamp / FAPESP, 2003 [1993].
- CARDINALETTI, A. Towards a cartography of subject positions. In: RIZZI, L. (org.) *The Structure of CP and IP*. New York: Oxford University Press. vol 2., 2004. p. 115-165.
- CARDINALETTI, A. Cross-linguistic variation in the syntax of subjects. In: PICALLO, M. C. (org.) *Linguistic Variation in the Minimalist Framework*, Oxford, Oxford University Press, 2014. p. 82-107.
- CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.
- CHOMSKY, N. *Barriers*. Cambridge: MIT Press.
- CHOMSKY, N. *The minimalist program*. Cambridge (MA): MIT Press, 1995.
- CINQUE, G.; RIZZI, L. The cartography of syntactic structures. In: BERND, H.; HEIKO, N. (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 2010. p. 65-78.
- COSTA, J.; DUARTE, I.; SILVA, C. Construções de redobro em português brasileiro: sujeitos tópicos vs. Soletração do traço de pessoa. *Leitura*, n. 33, p. 135-145, 2004. DOI: <https://doi.org/10.28998/2317-9945.200433.135-145>.
- DUARTE, M. E. L. *A perda do princípio “Evite Pronome” no português brasileiro*. 1995. 149 f. Tese (Doutorado em Estudos da Linguagem) - Instituto de Estudos da Linguagem, Unicamp, 1995.
- DUARTE, M. E. L.; KATO, M. A. Mudança paramétrica e orientação para o discurso. In: *Comunicação apresentada no XXIV Encontro da Associação Portuguesa de Linguística*, Universidade do Minho, Braga, 2008.
- FRASCARELLI, M.; HINTERHÖLZL, R. Types of Topic in German and Italian. In: SCHWABE, K.; WINKLER, S. (orgs.) *On Information Structure, Meaning and Form*. Benjamins, 2007. p. 87-116.
- FROTA, S.; OLIVEIRA, P.; CRUZ, M.; VIGÁRIO, M. P-ToBI: *Tools for the transcription of Portuguese prosody*. Lisboa: Laboratório de Fonética, CLUL/FLUL; 2015. Disponível em: <http://labfon.letras.ulisboa.pt/InAPoP/P-ToBI>. Acesso em: 25 nov. 2023.
- GASQUE DE SOUZA, K. *A duplicação de sujeito no português brasileiro*. 2021. 122 f. Dissertação (Mestrado em Estudos da Linguagem) - Faculdade de Letras, Universidade Federal do Rio Grande do Sul, 2021.
- GUJARATI, D. N.; PORTER, D. C. *Econometria básica*. 5. ed. Porto Alegre: AMGH, 2011.

- GUY, G. R.; ZILLES, A. *Sociolinguística quantitativa - Instrumental de análise*. São Paulo: Parábola Editorial, 2007.
- KRIECK, L. E. *As sentenças com duplicação do sujeito no português brasileiro: uma análise cartográfica*. 2022. 148 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade Federal de Santa Catarina, 2022.
- LADD, D. R. *Intonational Phonology*. Cambridge, Mass: CUP, 2008 [1996].
- MIYAGAWA, S. Unifying agreement and agreement-less languages. Proceedings of WAFL2. MIT *Working Papers in Linguistics*, 54, p. 47-66, 2007.
- MIYAGAWA, S. *Why agree? Why move? Unifying agreement-based and discourse configurational languages*. Cambridge: MIT Press, 2010.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to linear regression analysis*. John, Wiley and Sons, Inc., New York, 2006.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*: with a new foreword. 2nd ed. Berlin, New York: Mouton de Gruyter, 2007[1986].
- NICOLAU DE PAULA, M. *As construções de deslocamento à esquerda de sujeito no PB: um estudo em tempo real de curta duração*. 2012. 99 f. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2012.
- OLIVEIRA, A. J. Análise quantitative no estudo da variação linguística: noções de estatística e análise comparativa entre Varbrul e SPSS. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 17, n. 2, p. 93-119. 2009. DOI: <http://dx.doi.org/10.17851/2237-2083.17.2.93-119>.
- ORSINI, M. T. *As construções de tópico no Português do Brasil: uma análise sintático-discursiva e prosódica*. 2003. 197 f. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2003.
- ORSINI, M. T.; VASCO, S. L. Português do Brasil: língua de tópico e de sujeito. *Diadorm*, Rio de Janeiro, v. 2, p. 83-98, 2007. DOI: <https://doi.org/10.35520/diadorm.2007.v2noa3852>.
- PIERREHUMBERT, J. *The phonology and phonetics of english intonation*. PhD Thesis, MIT, 1980.
- PONTES, E. *O Tópico no Português do Brasil*. Campinas: Editora Pontes, 1987.
- QUAREZEMIN, S. A arquitetura da sentença no Português Brasileiro: considerações sobre Sujeito e Tópico. *Revista Letras*, Curitiba, n. 96, p. 196-218, 2017. DOI: <http://dx.doi.org/10.5380/rel.v96i0.51027>.
- QUAREZEMIN, S. Um novo olhar sobre as sentenças com redobro em Português Brasileiro. *Revista da ANPOLL*, v. 1, p. 52-63, 2019. DOI: <https://doi.org/10.18309/anp.v1i48.1253>.
- QUAREZEMIN, S. Brazilian double subjects and the sentence structure. In: PIRES DE OLIVEIRA, R.; EMMEL, I.; QUAREZEMIN, S. *Brazilian Portuguese, Syntax and Semantics: 20 years of Núcleo de Estudos Gramaticais*. John Benjamins Publishing Company, 2020.
- QUAREZEMIN, S.; TESCARI NETO, A. A propósito dos vinte e cinco anos do programa cartográfico no brasil: hierarquias cartográficas e explanação teórica. *Estudos Linguísticos e Literários*, Salvador, n. 77, p. 470–531, 2024. Disponível em: <https://periodicos.ufba.br/index.php/estudos/article/view/61694/32836>. Acesso em: 30 de agosto de 2024. DOI: <https://doi.org/10.9771/ell.voi77.61694>.

REZENDE DOS REIS, E. P. *O redobro do sujeito no Português Brasileiro e no Português Europeu: empirismo e formalismo*. 2023. 166 f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2023.

REZENDE DOS REIS, E. P.; DUARTE, M. E. L. . O redobro do sujeito no português brasileiro e no português europeu: empirismo e formalismo. *Estudos Linguísticos e Literários*, Salvador, v. 77, p. 361-387, 2024. DOI: <https://doi.org/10.9771/ell.v0i77.61681>.

REZENDE DOS REIS, E. P.; DUARTE, M. E. L. Uma proposta preliminar para a formalização do(s) redobro(s) do sujeito no PB. *Revista Linguística*. No prelo.

RIZZI, L. The Fine Structure of the Left Periphery. In: Haegeman, L. (ed.). *Elements of Grammar*. Kluwer, Dordrecht, 1997. p. 281-337.

ROSA-SILVA, F. Deslocamento de tópico contrastivo no português brasileiro: uma proposta semântico-pragmática. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 27, n. 2, p. 771-809, 2019. DOI: 10.17851/2237-2083.27.2.771-809.

ROSS, J. R. *Constraints on variables in syntax*. Ph.D. Thesis, Cambridge: MIT, 1967.

SERRA, C. R. *Realização e percepção de fronteiras prosódicas no Português do Brasil*. 2009. 241 f. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2009.

SILVA, S. M. *Construções de deslocamento à esquerda no gênero textual debate: uma análise na interface sintaxe-discurso-prosódia*. 2018. 132 f. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2018.

TENANI, L. E. *Domínios prosódicos do português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos*. 2002. 317 f. Tese (Doutorado em Linguística) - Instituto de Estudos da Linguagem, Universidade de Campinas, 2002.

TENANI, L. E. Fonologia Prosódica. In: DA HORA, D; MATZENAUER, C. L. (orgs.). *Fonologia, fonologias: uma introdução*. 1ed. São Paulo: Contexto, 2017. p. 109-123.

TESCARI NETO, A. *Sintaxe Gerativa: uma introdução à Cartografia Sintática*. 1. ed. Campinas: Editora da UNICAMP, 2021.

VASCO, S. *Construções de tópico na fala popular*. 2006. 216 f. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2006.

VIEIRA, S. R.; MOTA, M. A. C. da (Org.). *Corpus Concordância*. Rio de Janeiro: Faculdade de Letras/UFRJ. Disponível em: VIEIRA, S.R.; BRANDÃO, S.F. CORPORAPORT: Variedades do Português em análise. Rio de Janeiro: Faculdade de Letras/UFRJ. Disponível em: www.corporaport.letras.ufrj.br. Acesso em: 17 de jul. 2024.

YANO, C. T.; FERNANDES, F. R. Um estudo preliminar sobre a prosódia de construções com tópico e foco no português paulista. *Entrepalavras*, Fortaleza, v. 10, n. 1, p. 256-282, 2020. DOI: <http://dx.doi.org/10.22168/2237-6321-11724>.

Ferramentas analíticas para uma historiografia dos modelos sintáticos: rede taxonômica e glossário de metatermos da *Grammatica da lingua Portuguesa* (1540), de João de Barros

Analytical Tools for a Historiography of Syntactic Models: Taxonomic Network and Glossary of Metaterms from 'Grammatica da lingua Portuguesa' (1540) by João de Barros

Francisco Eduardo Vieira
Universidade Federal da Paraíba (UFPB)
João Pessoa | PB | BR
fevs@academico.ufpb.br
<https://orcid.org/0000-0001-5076-4488>

Resumo: Este artigo objetiva apresentar duas ferramentas para análise de modelos sintáticos na história da gramaticografia ocidental: a *rede taxonômica* e seu *glossário de metatermos*. Essas ferramentas são utilizadas na sistematização do modelo sintático contido na *Grammatica da língua Portuguesa* (1540), de João de Barros. A pesquisa alinha-se a reflexões sobre historiografia da gramaticografia e da terminografia (Swiggers, 2020; Gómez Asencio; Gonçalves, 2015; entre outros) e adere a certos princípios e procedimentos, como, por exemplo: a consideração do conteúdo dos metatermos na interface “conteúdo focal” e “conteúdo contrastivo”; e a análise dos metatermos em diferentes níveis de referência, dos mais “substanciais” aos mais “relacionais”. Os resultados mostram que a rede taxonômica registra graficamente e exemplifica as relações formais e lógicas entre os metatermos que estruturaram um modelo sintático, capturando sua macro-organização e revelando detalhes terminográficos. Já o glossário de metatermos permite ao leitor do presente um entendimento mais aprofundado de um modelo sintático do passado, bem como o acompanhamento das transformações das relações entre termos e conceitos. A partir da interpretação da rede e do glossário de Barros (1540), conclui-se, entre outros pontos, que: a ideia de “regimento” se impõe na obra não só pela possibilidade de obediência ao sistema



de casos latinos, mas também pelo seu valor ao eixo da descrição metalinguística; o gramático é sensível a construções complexas da língua, embora lhe falte aparato técnico e terminologia apropriada para sistematizar esses processos sintáticos.

Palavras-chave: Historiografia da Linguística. Gramaticografia. Terminografia. Sintaxe. João de Barros.

Abstract: This article aims to present two analytical tools for the analysis of syntactic models in the history of Western grammaticography: the *taxonomic network* and its *glossary of metaterms*. These tools are employed in systematizing the syntactic model found in the 'Grammatica da língua Portuguesa' (1540) by João de Barros. The research aligns with reflections on the historiography of grammaticography and terminography (Swiggers, 2020; Gómez Asencio; Gonçalves, 2015; among others) and adheres to certain principles and procedures, such as considering the content of metaterms within the 'focal content' and 'contrastive content' interface, and analyzing metaterms at different levels of reference, from the most 'substantial' to the most 'relational'. The results demonstrate that the taxonomic network graphically records and exemplifies the formal and logical relationships between metaterms that structure a syntactic model, capturing its macro-organization and revealing terminographic details. The glossary of metaterms allows present-day readers to gain a deeper understanding of a past syntactic model, as well as to track transformations in relationships between terms and concepts. From the interpretation of Barros's (1540) network and glossary, it is concluded, among other points, that the notion of 'regimento' asserts itself in the work not only due to the possibility of adherence to the Latin case system but also for its value to the metalinguistic description axis; the grammarian is sensitive to complex language constructions, although lacking technical apparatus and appropriate terminology to systematize these syntactic processes.

Keywords: Historiography of Linguistics. Grammaticography. Terminography. Syntax. João de Barros.

1 Por uma história dos modelos sintáticos: temas, campos e objetivos de pesquisa

Neste artigo, compartilho parte dos resultados de meu estágio de pós-doutoramento no Programa de Pós-Graduação em Letras (PPGL) da Universidade Federal do Paraná (UFPR), onde desenvolvi o projeto de pesquisa “*Historiografia da gramaticografia brasileira: modelos sintáticos*”, sob a supervisão do Prof. Dr. Carlos Alberto Faraco. Esse projeto se insere no macro-projeto de pesquisa “*Historiografia da Sintaxe no Brasil (HSB): teoria, norma e ensino*”, que coordeno no Programa de Pós-Graduação em Linguística (PROLING) e no Grupo de Pesquisa “*HGEL – Historiografia, Gramática e Ensino de Línguas*”¹, ambos sediados na Universidade Federal da Paraíba (UFPB).

O macroprojeto HSB tem como objetivo principal realizar uma historiografia de saberes, ideias, questões, teorias e modelos pedagógicos relacionados à dimensão sintática dos estudos gramaticais, filológicos e linguísticos no Brasil, abrangendo o período entre os séculos 19 e 21 (cf. Vieira, 2020a). Dentro desse amplo escopo, o projeto de estágio de pós-doutoramento concentra-se na análise de modelos sintáticos que foram elaborados e difundidos ao longo da história da gramaticografia da língua portuguesa. Busca-se compreender os processos de surgimento, desenvolvimento, recepção, contraposição, apagamento e continuidade desses modelos.

A base disciplinar deste artigo, construído a partir dessas pesquisas, é a Historiografia da Linguística, um campo de estudo que investiga a história dos conhecimentos e reflexões relacionados à linguagem e às línguas (Swiggers, 2009a, p. 68), respeitando os requisitos epistemológicos e metodológicos necessários à escrita de narrativas historiográficas adequadas (Koerner, 2020, p. 12). A tarefa de escrever uma história dos modelos sintáticos envolve, portanto, a seleção, ordenação, reconstrução, descrição e interpretação de fontes e conteúdos significativos e contextualmente situados. Além disso, requer a aplicação de princípios bem definidos e procedimentos básicos amplamente aceitos (Altman, 2012, p. 27-30).

Este trabalho, em particular, alinha-se às reflexões meta-historiográficas de trabalhos como Swiggers (2020, 2009b), Gómez Asencio, Montoro del Arco e Swiggers (2014) e Silva (2006) sobre *historiografia da gramaticografia*, definida como a escrita da história da técnica de compor gramáticas, ou seja, do ato de produzi-las. A gramaticografia representa o domínio de ação do gramático, englobando tarefas específicas, tomadas de decisões e implicações inerentes a esse processo; por seu turno, a historiografia da gramaticografia integra-se à Historiografia da Linguística, focando na comparação e no estabelecimento de relações entre as concepções de gramáticos e as soluções propostas para problemas gramaticais no curso da história.

Com efeito, a historiografia da gramaticografia pode ser vista como um domínio dentro da história das técnicas (de segmentação, classificação, combinação, correlação etc.), das noções (fonéticas, morfológicas, sintáticas, semânticas etc.) e das relações (formais, lógicas, funcionais etc.) utilizadas na análise, descrição, explicação e ensino das estruturas linguísticas. Grande parte dessas técnicas, noções e relações são compartilhadas pelas principais “linhagens” gramaticográficas de base greco-latina – as linhagens latinizada, racionalista e empirista (cf. Faraco; Vieira, 2021) –, apesar das diferenças conceptuais que as

¹ Espelho do grupo de pesquisa no Diretório dos Grupos de Pesquisa no Brasil (Plataforma Lattes/CNPq): <http://dgp.cnpq.br/dgp/espelhogrupo/6433198070413694>. Site do grupo: hgel.com.br. Acesso em: 13 set. 2024.

separam. Por essa razão, historiógrafos da gramaticografia, como Borges Neto (2022, p. 26), costumam utilizar o hiperônimo “gramática tradicional” para se referirem ao diversificado conjunto de gramáticas de diferentes línguas produzidas em mais de dois mil anos de atividades gramaticográficas no Ocidente:

Chama-se de *gramática tradicional* um tipo de estudo que tem origem nas reflexões filosóficas dirigidas às manifestações linguísticas pelos filósofos gregos. Por meio dessas reflexões filosóficas, foi se desenvolvendo ao longo do tempo um conjunto de noções e de uma terminologia correspondente que, por volta do início do século I a.C., resultou numa forma de “manual técnico” – a *Téchnē grammatiskē* – que tratava de aspectos das manifestações linguísticas e que se destinava, basicamente, a um uso escolar. Essas *Téchnai* foram reproduzidas pelos romanos – com o nome de *artes grammaticae* – e posteriormente, já na Idade Média, passaram a ser chamadas simplesmente de *gramáticas*.

Esse modelo de gramática, que vem dos gregos e dos latinos e que se espalha por toda a Europa a partir do século XVI, é seguido ainda hoje por nossas gramáticas escolares. É a esse modelo de gramática que vamos designar de *gramática tradicional* (ou, às vezes, de *gramática greco-latina*).

Ponto de vista semelhante apresento em Vieira (2020b, p. 89-95), quando argumento que a gramática tradicional, a despeito de seu caráter pedagógico e normativo, pode ser considerada uma “tradição de pesquisa” nos termos da filosofia da ciência de Larry Laudan (1941-2022). Laudan (2011, p. 113) define uma “tradição de pesquisa”, também chamada de “maxiteoria”, como “um conjunto de suposições acerca das entidades e dos processos de uma área de estudo e dos métodos adequados a serem utilizados para investigar os problemas e construir as teorias dessa área do saber”. Nesse sentido, tradições de pesquisa abrigam doutrinas conceitualmente relacionadas que as distinguem de outras tradições de pesquisa. Essas doutrinas configuram uma ontologia geral acerca da natureza dos objetos de cada tradição e se valem de coordenadas teórico-metodológicas também gerais, compatíveis com sua metafísica. Ainda de acordo com Laudan (2011, p. 112), as tradições de pesquisa são parcialmente constituídas por teorias específicas, que podem ser contemporâneas ou sucessoras umas das outras; costumam ter uma longa história que se estende por um significativo período de tempo; e apresentam diretrizes epistemológicas que, de maneira geral, especificam seus tipos de entidades fundamentais e certos modos de proceder, fornecendo coordenadas para desenvolvimento de modelos e métodos próprios.

A partir dessa perspectiva de “tradição de pesquisa”, a gramática tradicional pode ser compreendida como um conjunto de teorias ou modelos teóricos historicamente relacionados, desenvolvidos com base em diretrizes epistemológicas mais gerais². Embora essas teorias e modelos possam apresentar divergências quanto a técnicas, noções e relações específicas, eles compartilham claros vínculos metafísicos e metodológicos entre si. Essas diretrizes orientam a regulação normativa, representada pelo *eixo da norma-padrão*, e a descrição estrutural, representada pelo *eixo da análise metalinguística* de uma língua. Além disso, constituem uma ontologia que tanto sustenta quanto limita a ação do gramático, especificando os tipos de entidades fundamentais, bem como os princípios e procedimentos considerados legítimos para a gramatização:

² Para aprofundamento da discussão sobre a aplicabilidade do conceito de “tradição de pesquisa” (Laudan, 2011) à gramática tradicional, cf. Vieira (2020b).

Quadro 1 – Diretrizes epistemológicas da gramática tradicional

<p>1. A gramática tradicional busca estabelecer e ensinar um padrão linguístico idealizado por meio da prescrição de formas e construções supostamente corretas e legítimas.</p> <p>2. A gramática tradicional promove uma visão de língua invariável e imutável, interdita ou ignora a diversidade linguística e combate a suposta deterioração do padrão idealizado.</p> <p>3. A gramática tradicional privilegia a escrita literária do passado em detrimento de outras esferas discursivas e outras sincronias.</p> <p>4. A gramática tradicional considera a oração como a unidade máxima de análise e como a expressão de um juízo, dotada de sentido completo.</p> <p>5. A gramática tradicional utiliza um conjunto de categorias e conceitos oriundos da filosofia grega e da gramática greco-latina, adaptando-os às diversas línguas modernas.</p>	<p>Eixo da norma-padrão</p>
	<p>Eixo da análise metalinguística</p>

Fonte: Elaboração própria, com base em Vieira (2020b, p. 94).

No Quadro 1, o *eixo da norma-padrão* equivale ao campo da escrita e da fala “corretas”. Compreende a ortografia e as regras de acentuação gráfica, a ortoépia e a prosódia “elegante”, as flexões nominais e verbais, os paradigmas pronominais, a concordância, a regência e a colocação consideradas possíveis, as estruturas relativas padronizadas, o uso do acento indicativo de crase, as convenções de pontuação, entre outras regras e convenções próprias da escrita normatizada e da fala supostamente de prestígio. Já o *eixo da análise metalinguística* equivale ao domínio das técnicas de descrição e explicação das estruturas fonético-fonológicas, morfológicas e sintáticas da língua. Essas técnicas pressupõem um aparato categorial e conceitual específico, ou seja, uma terminologia gramatical relacionada a uma rede nocional. Note-se que esses dois eixos não são estanques, mas se articulam como duas engrenagens na composição da tradição gramatical: a prescrição da norma-padrão acontece ancorada em gestos de análise metalinguística que envolvem relações fonético-fonológicas, morfológicas e sintáticas entre os elementos da gramática da língua, por sua vez nomeados a partir de uma taxonomia específica, de origem greco-latina.

Os modelos sintáticos que caracterizam a história da gramaticografia de língua portuguesa no Brasil a partir do século 19, em sua maioria, foram elaborados e difundidos em consonância com essas cinco diretrizes epistemológicas do Quadro 1. O primeiro desses modelos é encontrado na “primeira tentativa de descrição sistemática do português edificada por um brasileiro” (Cavaliere, 2014, p. 60): o *Epitome da Grammatica da Lingua Portugueza* (1806), obra de autoria do lexicólogo, gramático e tradutor Antonio de Moraes Silva (1755-1824), natural do Rio de Janeiro. Embora tenha sido publicado em Lisboa, o *Epitome* foi concluído em Pernambuco³, onde Moraes Silva viveu desde 1794 até seu falecimento.

De acordo com Borges Neto (2022, p. 493), modelos sintáticos em geral, da Antiguidade Clássica à contemporaneidade, são permeados por dois tipos de relações: as *relações formais* e as *relações lógicas*. As primeiras, relacionadas à concatenação de palavras na construção de unidades maiores, costumam circunscrever as marcas formais de concordância e regência presentes nas palavras. Elas integram o que denomino *eixo da norma-padrão*. Já as relações lógicas, vinculadas ao domínio da unidade proposicional, transfor-

³ Lê-se, ao fim da exposição gramatical: “Acabou se este Epitome da Grammatica Portugueza no Engenho novo da Moribeca em Pernambuco, aos 15. de Julho de 1802” (Moraes Silva, 1806, p. 163).

mam uma cadeia de palavras em uma expressão de sentido completo. Elas correspondem ao que entendo por *eixo da análise metalinguística*. Ao problematizar o desenvolvimento desses modelos sintáticos, Colombat, Fournier e Puech (2017, p. 128-133) afirmam que, embora os elementos para o estabelecimento das relações lógicas tenham sido fornecidos na filosofia grega (cf. Aristóteles, 1995), foi necessário esperar pelo modelo sintático de Port-Royal (cf. Arnauld; Lancelot, 1660) para que uma análise lógica da proposição, com *sujeito*, *cópula* (*verbo*) e *atributo*, fosse efetivamente incorporada à gramaticografia ocidental. Como se sabe, esse modelo fundamentou a sintaxe das gramáticas de linhagem racionalista – como o *Epitome* de Moraes Silva (1806) – e predominou em boa parte da gramaticografia portuguesa do século 19, tanto em Portugal quanto no Brasil.

É possível afirmar, portanto, que o primeiro modelo sintático apresentado em uma gramática brasileira de língua portuguesa se estrutura a partir de dois eixos, consoante as diretrizes epistemológicas da gramática tradicional (ver Quadro 1). O primeiro eixo – o da norma-padrão – aborda, por um lado, as “regras, que ensinão a mostrar as connexões entre os nomes, e os adjetivos, e os verbos”, isto é, a “Sintaxe de Concordancia”; e, por outro lado, as “regras, que ensinão a mostrar as relações entre os nomes, por meyo das preposições, e casos, ou da collocação”, isto é, a “Sintaxe de Regencia” (Moraes Silva, 1806, p. 84-86). Além disso, responde pelas “semelhanças de incorrecção, ou *Figuras*”, ou seja, pelos casos em que “a incorrecção é apparente, e dá uma nova figura, ou apparencia á composição, que por isso se diz *figurada*” (Moraes Silva, 1806, p. 103). O segundo eixo – o da análise metalinguística – equivale à “boa composição das partes da oração entre si”, da qual “resulta a *sentença*, ou sentido perfeito”, cujas partes são “sujeito”, “atributo”, “verbo” e, eventualmente, “circunstancias de lugar, tempo, modo, instrumento, fim, &c.” (Moraes Silva, 1806, p. 82-83).

O modelo sintático de Moraes Silva (1806) representa, sem dúvida, um marco histórico que assinala o início do pensamento gramaticográfico brasileiro. A abrangência desse modelo, no entanto, não se restringe a um momento específico (o começo do século 19); ele se insere, como diria Auroux (2014, p. 12), na “temporalidade ramificada da constituição cotidiana do saber”, apontando tanto para um “horizonte de retrospecção” (memória) quanto para um “horizonte de projeção” (futuro). As instâncias desse modelo sintático não apagam seu passado, mas sim o selecionam, organizam, adaptam, reconstruem e idealizam, ao mesmo tempo que antecipam seu futuro: “sem passado e sem projeto, não há saber” (Auroux, 2014, p. 12). Com efeito, o conhecimento sintático em Moraes Silva (1806) é fonte dialógica de nosso então porvir terminográfico, que culminaria na Nomenclatura Gramatical Brasileira (NGB) em 1959, seguida estritamente pelas gramáticas tradicionais subsequentes e parcialmente pelas gramáticas de linguistas publicadas no século 21. O Quadro 2 exemplifica esse horizonte de projeção, ao comparar os principais metatermos do modelo sintático do *Epitome* (1806) com os metatermos correspondentes, em forma e/ou conteúdo, da NGB (1959) e de uma recente gramática de linguistas contemporâneos, a *Gramática do português brasileiro escrito* (Vieira; Faraco, 2023). A semelhança entre os três projetos terminográficos é notável:

Quadro 2 – Horizonte de projeção dos principais metatermos da sintaxe de Moraes Silva (1806)

Instrumentos gramaticográficos			
	Moraes Silva (1806)	NGB (1959)	Vieira; Faraco (2023)
Metatermos	concordância	concordância	concordância
	regência	regência	regência
	figura	figura	-
	sujeito	sujeito	sujeito
	verbo	verbo	verbo
	atributo	predicativo do sujeito complemento verbal	complemento verbal
	circunstância	adjunto adverbial	adjunto adverbial
	oração sentença	oração período	oração período

Fonte: Elaboração própria.

Quanto ao horizonte de retrospecção das ideias sobre sintaxe em Moraes Silva (1806), estas ecoam indubitavelmente os três primeiros séculos da gramaticografia portuguesa (séculos 16 a 18), englobando não apenas relações formais de concordância e regência, mas também relações lógicas que envolvem as partes e o mecanismo organizacional da oração perfeita, de sentido completo.

Colombat, Fournier e Puech (2017, p. 129) estão corretos ao afirmar que, na história da sintaxe, a análise lógica da proposição foi de fato sistematizada nas gramáticas das línguas modernas europeias somente a partir do século 17, com os racionalistas franceses de Port-Royal. É notório que, até então, os modelos sintáticos ocupavam nas gramáticas espaços pouco significativos em comparação com os níveis dos sons/letras e palavras. O foco predominante recaía no componente morfológico-semântico das formas gramaticais, relegando o componente sintático, num sentido amplo, à categorização e análise das partes da oração (com base em critérios raramente sintáticos) e, num sentido estrito, como já dito, às regras de concordância e regência, bem como à listagem de solecismos (figuras de construção que abarcam problemas normativos).

Todavia, é possível identificar modelos sintáticos envolvendo relações propriamente lógicas em pelo menos dois importantes instrumentos da gramaticografia greco-latina anteriores aos racionalistas de Port-Royal. O primeiro deles é o tratado grego *Syntaksi* (Σύνταξη), escrito pelo gramático alexandrino Apolônio Díscolo no século 2 EC⁴, no qual já se anteveem ideias linguística como “sujeito” e “objeto”⁵. O segundo instrumento é a obra latina *Institutiones*

⁴ EC é a abreviatura de “Era Comum”, expressão alternativa para se referir ao período “depois de Cristo” (d.C.).

⁵ Segundo Botas (1987, p. 40-41), “Apolonio tiene muy claras las diversas funciones sintácticas o semántico-sintácticas; así, las de sujeto-objeto se expresan de muy diversas maneras [...] el objeto indirecto es, asimismo, considerado [...] En Apolonio la consideración de caso recto-caso oblicuo no es morfológica sino sintáctica: el caso «recto» es el que está «ordenado» o está em «ordem coincidente» con la persona verbal [...] es, por tanto, el caso del sujeto y sus aposiciones –el vocativo–, frente al oblicuo, que es el «desviado» o no coincidente con la persona del verbo [...] Apolonio intuye también la noción de complemento cuando habla de los acompañantes de los verbos [e aborda] los conceptos de actividad-pasividad y transitividad-intransitividad [...]. Em português: “Apolônio tem muito claras as diversas funções sintáticas ou semântico-sintáticas; assim, as de sujeito-objeto são expressas de diversas maneiras [...] o objeto indireto é, do mesmo modo, considerado [...] Em Apolônio, a consideração de caso reto-caso oblíquo não é morfológica, mas sintática: o caso ‘reto’ é aquele que está ‘ordenado’ ou ‘em ordem coincidente’ com a pessoa verbal [...] é, portanto, o caso do sujeito e suas aposições –o vocativo–, frente ao oblíquo, que é o ‘desviado’ ou não coincidente com a pessoa do verbo [...] Apolônio intui também a noção de complemento quando fala dos acompanhantes dos verbos [e aborda] os conceitos de atividade-pasividade e transitividade-intransitividade [...]”.

Grammaticae, escrita pelo gramático Prisciano de Cesareia no século 6 EC. Seu Livro XVII, o primeiro de um par dedicado à sintaxe, objetiva examinar o modo habitual de se organizar as palavras na construção de enunciados completos⁶, o que deságua, no correr dos capítulos, por exemplo, nas ideias de “transitividade” e “intransitividade” (cf. Priscien, 2010, p. 63-103; p. 155-169). Além disso, os gramáticos especulativos (modistas) da Idade Média (1270-1330), fundamentados na correspondência especular entre mundo/intelecto/linguagem e na lógica aristotélica, consideravam o par “suposto-aposto” (*suppositum-appositum*) essencial à construção sintática bem formada: “O suposto (um nominal) faz referência a uma substância no mundo, por isso, antecede logicamente o seu aposto, que se refere a um acidente, ou predicado, dessa substância” (Beccari, 2019, p. 51). Os modistas também analisaram a frase com base nas ideias de “dependente” (termo que deve/pode exigir a presença de outro) e “terminante” (termo que satisfaz tal exigência), reconhecendo, assim, a existência de relações sintáticas diferentes das relações formais de concordância (Robins, 1979, p. 65).

Esse conjunto de saberes, conhecimentos ou ideias sobre sintaxe envolvendo relações formais (práticas, normativas, do eixo da *norma-padrão*) e relações lógicas (análíticas, descriptivas, do eixo da *análise metalinguística*), apresentado em espaços múltiplos e em uma temporalidade descontínua, deixou marcas na linhagem latinizada da gramaticografia portuguesa. É verdade que o pioneirismo de Fernão d’Oliveira (1507-1581) em sua *Grammatica da lingoagem portuguesa* (1536) não nos oferece sequer um prenúncio de modelo sintático inferível. Embora ele faça menção à ideia de “côstruïção”, entendida como a “côpoisão ou concerto que as partes ou diçoës da nossa lingua tẽ antre si como em qualqr outra lingua” (Oliveira, 1871, p. 117), praticamente não há saberes sintáticos desenvolvidos e sistematizados na obra. Ao assunto só é dado espaço no capítulo XLIX – o penúltimo da gramática, pouco maior que uma página. Nesse capítulo, lê-se que a sintaxe deveria ser tratada com fôlego em obra posterior⁷, que, por algum motivo, nunca veio a ser escrita (ou talvez tenha se perdido).

Em contrapartida, instrumentos linguísticos portugueses (gramáticas e tratados ortográficos) subsequentes à *Grammatica* de Oliveira, publicados nos séculos 16 a 18, apresentam, em graus de explicitude diferenciados, modelos sintáticos diversos, que abrangem relações tanto formais (majoritárias) quanto lógicas (minoritárias), sustentadas pelas diretrizes epistemológicas da gramática tradicional. Embora parte dessas obras já tenha sido analisada em diferentes recortes e com outros propósitos, os modelos sintáticos que elas contêm ainda não foram minuciosamente descritos, sistematizados, interpretados e comparados entre si, a serviço de uma historiografia que os considere no horizonte de retrospecção imediato dos modelos sintáticos brasileiros que vieram à lume a partir do século 19 e que se projetam, ainda hoje, nas pesquisas linguísticas, nos instrumentos gramaticais e nas práticas escolares de ensino de gramática:

nado’ ou está em ‘ordem coincidente’ com a pessoa verbal [...]; é, portanto, o caso do sujeito e suas aposições – o vocativo –, em oposição ao oblíquo, que é o ‘desviado’ ou não coincidente com a pessoa do verbo [...] Apolônio também intui a noção de complemento quando fala dos acompanhantes dos verbos [e aborda] os conceitos de atividade-passividade e transitividade-intransitividade [...]” (tradução nossa).

⁶ A sintaxe de Prisciano é assim introduzida no Livro XVII das *Institutiones Grammaticae*: “Dans ce qui a été antérieurement exposé, nous avons donc traité des mots réalisés individuellement, comme le réclamait leur logique propre. Nous allons à présent parler de la façon habituelle de les agencer pour construire un énoncé complet [...]” (Prisciano, 2010, p. 63). Em português: “No que foi anteriormente exposto, tratamos das palavras realizadas individualmente, conforme exigia sua própria lógica. Vamos agora falar da maneira habitual de organizá-las para construir um enunciado completo [...]” (tradução nossa).

⁷ “Nesta derradeira parte ũ he da cõstruïção ou cõpoisão da lingua naõ dizemos mais porq temos começada húa obra em ũ particularmête e cõ mais comprimento falamos della.” (Oliveira, 1871, p. 118).

Quadro 3 – Instrumentos linguísticos portugueses (séculos 16 a 18) no horizonte de retrospecção dos modelos sintáticos brasileiros

Ano	Autor	Título	Cidade	Editora	Páginas
1540 1 ^a edição	João de Barros (1496-1570/71)	Grammatica da lingua Portuguesa Dialogo em louvor da nossa lingvagem	Lisboa	Olyssippone. Apud Lodouicūm Rotorigiū	120 (60 não p.)
1619 1 ^a edição	Amaro de Roboredo (1580/85-1653)	Methodo grammatical para todas as lingvas	Lisboa	Pedro Craesbeeck	241
1721 1 ^a edição	Caetano Maldonado da Gama ⁸	Regras da lingua Portugueza, eſpelho da lingua Latina, Ou disposiçām Para facilitar o enſino de lingua Latina pelas regras da Portugueza	Lisboa	Officina de Mathias Pereyra da Sylva & Joaõ Antunes Pedrozo	228
1725 [1721] 2 ^a impressão	Jeronymo Contador de Argote (1676-1749)	Regras da lingua Portugueza, eſpelho da lingua Latina, Ou disposiçāo para facilitar o enſino de lingua Latina pelas regras da Portugueza	Lisboa	Officina da Musica	356
1788 [1769] 6 ^a impressão	Joaõ Pinheiro Freire da Cunha (1738-1811)	Breve tratado da orthografia para os que naõ frequentaraõ os estudos [...] ⁹	Lisboa	Officina de Antonio Gomes	iv + 202
1770 1 ^a edição	Antonio José dos Reis Lobato (1721-1804)	Arte da grammatica da lingua portugueza [...] ¹⁰	Lisboa	Regia Officina Typografica	xlviii + 253
1797 [1770] 4 ^a impressão	Antonio José dos Reis Lobato (1721-1804)	Arte da grammatica da lingua portugueza [...] ¹¹	Lisboa	Regia Officina Typografica	xxxi + 229 + 1ps/n

⁸ Padre Caetano Maldonado da Gama é pseudônimo do padre Dom Jeronymo Contador de Argote. Segundo Kemmler (2012, p. 93), “Caetano” identifica o autor como Clérigo Regular (teatino); “Maldonado” é o apelido principal do avô paterno; e “da Gama” é o primeiro apelido do lado da avó materna.

⁹ Título completo: *Breve tratado da orthografia para os que naõ frequentaraõ os estudos, ou dialogos Sobre as mais principiaes Regras da Orthografia uteis para o Povo menos instruido, e para os que naõ tendo frequentado as Aulas, se achaõ hoje empregados nos Eſcritorios publicos, e dezejaõ acertar na praxe fẽm grande multiplicidade de regras, que naõ lhes ſão facieis de comprehendere, e muito mais proveitózios aos Meninos, que frequentaaõ as Eſcolas.* A 1^a edição da obra, indisponível no acervo de fontes historiográficas do grupo de pesquisa HGEL/UFPB, foi assinada pelo autor sob o pseudônimo Domingos Dionysio Duarte Daniel.

¹⁰ Título completo: *Arte da grammatica da lingua portugueza, composta, e offerecida ao III.^{mo} e Exc. ^{mo} Senhor Sebastião José de Carvalho e Mello Conde de Oeyras Ministro, e Secretario de Estado de Sua Mageſtade Fidelissima, da Repartição dos Negocios do Reino, Alcaide Mór da Cidade de Lamego, e Senhor Donatario das Villas de Oeyras, Pombal, Carvalho, e Cercofa, e dos Reguengos, e Direitos Reaes da De Oeyras, e de Apar de Oeyras, Commendador das Commendas de Santa Marinha de Mata Lobos, e de S. Miguel das tres Minas na Ordem de Chriſto, &c.*

¹¹ Título completo: *Arte da grammatica da lingua portugueza, composta, e offerecida ao III.^{mo} e Exc. ^{mo} Senhor Sebastião José de Carvalho e Mello, Marquez de Pombal, Ministro, e Secretario de Estado de Sua Mageſtade Fidelissima da Repartição dos Negocios do Reino, &c.*

1783 1 ^a edição	Bernardo de Lima e Mélo Bacellar (1736-1787)	Grammatica philosophica, e orthographia racional da Lingua Portugueza; Para se pronunciarem, e escreverem com acerto os vocabulos deste idiôma	Lisboa	Offic. de Simão Thaddeo Ferreira	196 + 1ps/n
1786 1 ^a edição	Francisca de Chantal Álvares ¹² (1745-?)	Breve compendio da gramatica portugueza para o uso Das Meninas que se educaõ no Mosteiro da Vizitaõ de Lisboa	Lisboa	Officina de Antonio Rodrigues Galhardo	51 + 3ps/n
1792 1 ^a edição	Joaõ Joaquim Casimiro (?-?) ¹³	Methodo grammatical resumido da lingua portugueza	Porto	Offic. de Antonio Alvarez Ribeiro	127
1799 1 ^a edição	Pedro José de Figueiredo (1762-1826)	Arte da grammatica portugueza ordenada em methodo breve, facil, e claro	Lisboa	Regia Officina Typografica	113 + 3ps/n
1799 1 ^a edição	Pedro José da Fonseca (1736-1816)	Rudimentos da grammatica portugueza, Cómmodos á instrucção da Mocidade, em confirmados com selectos exemplos de bons Autores	Lisboa	Off. de Simão Thaddeo Ferreira	xv + 353

Fonte: Elaboração própria.

O mapeamento epi-historiográfico¹⁴ apresentado no Quadro 3 totaliza onze gramáticas e um tratado ortográfico correspondentes às edições dos instrumentos linguísticos portugueses que atendem aos seguintes critérios: i) estão disponíveis para análise no acervo de fontes historiográficas do grupo de pesquisa HGEL/UFPB¹⁵; ii) foram publicadas nos séculos 16 a 18; iii) apresentam saberes (ideias/conhecimentos) sintáticos de caráter formal e/ou lógico, em graus variados de explicitude e sistematização. Para atender ao critério (iii), um conjunto maior de fontes foi preliminarmente analisado, resultando na exclusão de algumas obras desse mapeamento. Entre as obras excluídas, destacam-se a já mencionada *Grammatica da lingoagem portuguesa* (1536), de Fernão d'Oliveira; a *Grammatica da lingua portuguesa com os mandamentos da Santa mádre igreja* (1539), de João de Barros; as *Regras que ensinam a maneira de escrever e a ortografia da língua portuguesa* (1574), de Pêro de Magalhães de Gândavo (1540-

¹² A obra somente é assinada em seu prefácio, com as iniciais F. C. – abreviação para Francisca de Chantal. Segundo Kemmler, Assunção e Fernandes (2010, p. 390), a irmã Francisca de Chantal Álvares foi uma freira do Mosteiro da Visitação, em Lisboa, que, antes de se tornar religiosa visitandina, se chamava Ana Inácia do Coração de Jesus.

¹³ Segundo Schäfer-Prieß (2019, p. 42) e Boliqueime (2021, p. 84), quase nada se sabe sobre a vida de João Joaquim Casimiro, além do fato de o gramático ter nascido no Porto, cidade em que também foi mestre de primeiras letras.

¹⁴ De acordo com Swiggers (2010a, p. 5; 2013, p. 4), a epi-historiografia é uma atividade que integra o material documental produzido por historiógrafos e apoia sua prática descritivo-interpretativa. Atividades de edição e tradução de fontes primárias e atividades de documentação bibliográfica, como a lista de instrumentos linguísticos apresentada no Quadro 3, são exemplos de epi-historiografia.

¹⁵ Desde a fundação do grupo de pesquisa HGEL/UFPB, em 2017, temos dedicado esforços contínuos à construção e organização deste acervo. Em constante expansão, a coleção visa abranger, entre outros conjuntos de fontes, o panorama gramaticográfico da língua portuguesa desde o século 16 até o século 21. Até dezembro de 2023, catalogamos aproximadamente 200 instrumentos linguísticos, predominantemente gramáticas, disponíveis em formato digital ou em edições impressas.

1579); a *Orthographia da lingoa Portuguesa* (1576), de Duarte Nunez de Lião (1530-1608); e a *Orthographia da Lingua Portugueza* (1736), de Luis Caetano de Lima (1671-1757).

Minha hipótese – cuja averiguação não se encerra neste artigo, mas se estende a trabalhos subsequentes – é que uma análise descritiva, interpretativa e comparativa dos modelos sintáticos contidos nesses doze instrumentos, se configurada a serviço (e por meio) da elaboração de redes taxonômicas e glossários de metatermos, poderá caracterizar o horizonte de retrospecção imediato dos primeiros modelos sintáticos brasileiros desenvolvidos a partir do século 19, bem como possibilitar uma compreensão mais aprofundada da produção, circulação e evolução desses modelos, os quais ainda não foram suficientemente detalhados, em termos de continuidades e rupturas, sob uma perspectiva panorâmica e serial¹⁶.

A possibilidade de desenvolver uma historiografia panorâmica e serial desses modelos sintáticos portugueses, do século 16 ao 18, ancora-se na metodologia da História Serial (cf. Barros, 2012). Essa abordagem, salvo engano, foi introduzida na historiografia da gramaticografia brasileira por Polachini (2018), ao mapear 200 gramáticas brasileiras do português publicadas no século 19 e relacionar dados como título, subtítulo, autor, ano de publicação, local e casa impressora. As fontes na metodologia da História Serial não são tratadas de forma isolada, analisadas em seus próprios limites, mas consideradas constituintes de uma grande cadeia temporalizada e especializada de fontes de mesmo tipo, tornando-as comparáveis entre si. O propósito da comparação entre as fontes de um conjunto serial é identificar tanto as continuidades quanto as variações e mudanças nos modelos sintáticos dentro de um período específico. É relevante afirmar que a análise serial não equivale à análise quantitativa, pois seu interesse está nas recorrências (ou na falta delas) e não nos valores numéricos. Por essa razão, ela pode prescindir de uma abordagem estatística computacional.

Considerando o raciocínio estabelecido até aqui, este artigo tem dois claros objetivos que se relacionam e se inserem nesse projeto historiográfico mais amplo sobre os modelos sintáticos da gramaticografia brasileira a partir do século 19. O primeiro objetivo, de natureza meta-históriográfica, é apresentar duas ferramentas inéditas e significativas para análise, reconstrução, sistematização e comparação de modelos sintáticos explícitos ou subjacentes na história da gramaticografia ocidental: a *rede taxonômica* e seu *glossário de metatermos*.

Por meio da rede taxonômica, o historiógrafo da gramaticografia poderá registrar graficamente e exemplificar as relações formais e lógicas entre os metatermos que estruturam um determinado modelo sintático, capturando sua macro-organização e revelando detalhes terminográficos, não raro despercebidos, envolvendo hierarquização, classificação e categorização de metatermos. A natureza desse registro é simultaneamente panorâmica e sinóptica, pois tanto proporciona uma visão ampla e abrangente das ideias sintáticas de um autor ou obra quanto destaca minudências e sintetiza informações para tornar essas ideias mais compreensíveis e acessíveis, além de possibilitar a comparação com outras ideias sobre sintaxe no curso da história da gramaticografia.

¹⁶ Evidentemente, há produtivas historiografias da gramaticografia brasileira que, no domínio da sintaxe, abordam um autor, obra ou período específico (cf., por exemplo, Dias, 2008; Polachini, 2010-2011, 2016; Polachini; Danna, 2016), ou se concentram em uma categoria ou fenômeno gramatical (cf., por exemplo, Leite, 2013; Moura Neves, 2014; Polachini, 2018; Ferreira, 2021). Também devem ser mencionadas as historiografias panorâmicas da gramaticografia, que, devido à natureza do produto, costumam apresentar considerações importantes sobre os estudos sintáticos em diferentes momentos da gramaticografia brasileira (cf., por exemplo, Vieira, 2018, 2020a, 2020b; Borges Neto, 2022; Cavaliere, 2014, 2022). O conjunto desses trabalhos, entretanto, não invalida nosso ponto de vista.

Já a partir do glossário dos metatermos da rede taxonômica, o historiográfico lida com os significados dos metatermos – isto é, com a “dimensão conceitual da terminologia” (Swiggers, 2009b, p. 14) – e opera uma espécie de “adequação” (Koerner, 1996a, p. 60) das ideias sintáticas do autor ou obra do passado a um suposto repertório gramatical comum ao leitor do presente. Isso permite a esse leitor um entendimento mais aprofundado e preciso do modelo sintático sob análise, bem como o acompanhamento das transformações das relações entre um metatermo e um conceito sintático na gramaticografia de um autor, de uma linhagem ou de um período específico da história (as duas últimas possibilidades a partir de um conjunto serial de fontes). Sabe-se que a relação “metatermo-conceito” – nem sempre unívoca ou transparente – pode mudar significativamente de um instrumento para outro e de uma época para outra. O glossário de metatermos, pois, é uma importante ferramenta para sistematizar a terminologia gramatical, da qual a historiografia da gramaticografia e as ciências da linguagem em geral podem se beneficiar.

O segundo objetivo, de natureza historiográfica, é expor, por meio de uma rede taxonômica e de seu glossário de metatermos, as ideias sobre sintaxe contidas no primeiro instrumento linguístico listado no mapeamento epi-historiográfico do Quadro 3: a *Grammatica da língua Portuguesa* (1540), de João de Barros. Ambas as ferramentas, portanto, são utilizadas na fase de exposição dos resultados do trabalho historiográfico (Swiggers, 2009a, p. 71), ou seja, na demonstração dos resultados da análise, reconstrução e sistematização do primeiro modelo sintático elaborado na gramaticografia portuguesa, particularmente na obra que deu início às gramáticas tradicionais/normativas luso-brasileiras de base greco-latina (Barros, 2011, p. 295) e cuja terminologia continua viva nas gramáticas contemporâneas do português (Gómez Asencio; Gonçalves, 2015, p. 111).

2 Bases epistemológicas para análise, reconstrução e sistematização de modelos sintáticos

Tendo situado este trabalho, de modo amplo, na Historiografia da Linguística e, de modo específico, na historiografia da gramaticografia, convém agora explicitar algumas noções e procedimentos teórico-metodológicos desses campos, os quais sustentam as análises e os produtos (meta-)historiográficos desenvolvidos e apresentados neste artigo.

2.1 Ideia linguística, modelo sintático, terminografia linguística e terminologia sintática

Em primeiro lugar, retomo a noção de que o objeto da Historiografia da Linguística não é a linguagem e as línguas propriamente ditas, mas o curso evolutivo do conhecimento (ou saber) sobre a linguagem e as línguas. Dito de outro modo, esse campo disciplinar se interessa pelo devir histórico das *ideias linguísticas*, em termos de aquisição, elaboração, formulação, circulação, difusão, recepção, influência, longevidade, apagamento, entre outras nuances.

Nessa perspectiva, “conhecimento”, “saber” e “ideia linguística” são expressões equivalentes e bastante amplas. Por exemplo, uma teoria, uma argumentação, uma asserção,

uma observação, uma pressuposição, uma prescrição ou um preconceito envolvendo linguagem ou línguas são considerados ideias linguísticas, ou seja, fazem parte do conjunto de conhecimentos ou saberes linguísticos. Essas expressões também referenciam “reflexões, especulações, descrições, análises, regras normativas sobre a linguagem e as línguas que foram elaboradas séculos antes de se poder efetivamente falar de uma linguística” (Batista, 2020, p. 18). A circunscrição e a análise de uma ideia linguística só podem ser realizadas levando em conta sua natureza historicamente contextualizada e a relação, em termos de continuidade e ruptura, entre a ideia mapeada e outras ideias linguísticas e demandas socioculturais de diferentes épocas.

Um modelo sintático contido numa gramática normativa ou, eventualmente, num tratado ortográfico é, portanto, uma ideia linguística ou, em último termo, um conjunto de ideias linguísticas relacionadas. Ele abrange a macro e a microorganização do que se entende por sintaxe no instrumento linguístico, o que inclui aspectos metalinguísticos e conceituais relativos a *categorias, propriedades, prescrições e relações*. Na história da gramatização portuguesa, os modelos sintáticos acompanham as diretrizes epistemológicas da gramática tradicional (ver Quadro 1). Contudo, variações, que podem se manifestar de maneira mais ou menos pronunciada – no que diz respeito à concepção de linguagem, língua, gramática e sintaxe, à organização da exposição, às descrições estruturais, às técnicas de análise, ao estabelecimento de regras normativas, às classificações e terminologias utilizadas nas obras –, resultam em variações no modelo sintático, mesmo quando se consideram instrumentos linguísticos pertencentes a um mesmo período ou a uma mesma linhagem gramaticográfica.

Dado que um modelo sintático é um conjunto de ideias linguísticas do domínio da gramaticografia, justifica-se sua análise a partir dos três componentes para exame interno de fontes gramaticográficas apresentados em Gómez Asencio, Montoro del Arco e Swiggers (2014, p. 282-283) e adaptados a seguir.

O primeiro deles é o componente *teórico*. Quando se trata da análise de modelos sintáticos, esse componente deve abranger: a noção de língua e gramática em que se insere o modelo; a noção de sintaxe e a macro-organização do modelo; as categorias, subcategorias e propriedades utilizadas na abordagem da sintaxe; e as relações entre esses elementos.

O segundo componente é o *descritivo-normativo*, que, no contexto analítico em discussão, vai corresponder à apreciação sociolinguística da sintaxe gramatizada no tocante à defesa de estabilidades e/ou incorporação de mudanças. O olhar aqui deve se voltar a: construções sintáticas interditadas e/ou legitimadas; frases que explicam e/ou exemplificam descrições, prescrições e proscrições; e paradigmas descritivo-normativos presentes no modelo sintático.

Por fim, o terceiro componente para o exame de modelos sintáticos é o componente *pedagógico*, que visa envolver: a disposição didática do conhecimento sobre sintaxe; as estratégias de aproximação do conselente, como adaptação de conteúdos e explicações; o grau de exemplificação; e os eventuais exercícios.

Embora não sejam descartadas considerações gerais sobre a natureza pedagógica de Barros (1540), seu modelo sintático será analisado, fundamentalmente, a partir dos componentes teórico e descritivo-normativo. Esse recorte encaminha o estudo na direção de uma *historiografia da terminografia linguística*, tendo como objeto uma *terminologia sintática*. Esta última é compreendida como um conjunto de entidades lexicais e construções complexas (metatermos) que estão em uso como formas designativas, para fins técnicos, em referência aos objetos (entidades e fatos) que constituem um modelo sintático. A análise terminogrâ-

fica em questão é do tipo “monográfica/isotópica”, nos termos de Swiggers (2009b, p. 24), pois se concentra na terminologia sintática de uma obra específica de um autor particularizado.

Ciente de que a historiografia da terminografia linguística envolve compilação, sistematização, leitura crítica e comparação de terminologias (Swiggers, 2006, p. 15), em trabalhos futuros, de posse dos resultados apresentados neste artigo, pretendo desenvolver uma análise terminográfica do tipo “comparativa/contrastiva” (Swiggers, 2009b, p. 24), abrangendo o cotejo e a contraposição entre o modelo sintático de Barros (1540) e os modelos sintáticos dos outros instrumentos linguísticos listados no Quadro 3, com o propósito de elaborar uma síntese da sintaxe nos três primeiros séculos de gramaticografia portuguesa, prenúncio da gramaticografia brasileira. Trata-se de algo não efetivamente realizado até o momento: uma história longitudinal e acumulativa da terminologia sintática luso-brasileira e um glossário substancial dessa terminologia.

2.2 O problema da metalinguagem, dimensões de análise e fases da pesquisa

Entre as dimensões de uma análise terminográfica, a metalinguagem talvez seja a mais evidente e complexa. Partindo de um cenário meta-historiográfico mais amplo que o escopo estritamente terminográfico, Koerner (1996b, p. 98) argumenta que o historiógrafo, ao lidar com ideias linguísticas do passado e tentar torná-las acessíveis ao leitor do presente, enfrenta problemas terminológicos por agir num nível meta-metalinguístico, os quais podem levar a distorções das ideias originais. Como resposta a essa questão, que ele identifica como o “problema da metalinguagem”, Koerner (1996a, p. 60) desenvolveu três princípios para a prática historiográfica. Esses princípios, que se tornaram amplamente reconhecidos e utilizados na historiografia da linguística brasileira, são sumarizados na sequência e considerados na análise do modelo sintático de Barros (1540).

O primeiro princípio é o da *contextualização*. Trata-se do estabelecimento do “clima de opinião” (Becker, 1932, p. 1-31), da atmosfera intelectual, do cenário cultural e epistemológico geral da época em que a *Grammatica da lingua Portuguesa* foi publicada. Talvez hoje seja um “truísmo historiográfico” dizer que fatores contextuais causam impacto nas ideias linguísticas de um autor ou obra (Koerner, 2014a, p. 12), ou que as ideias linguísticas não se desenvolvem no vazio, desvinculadas de outras ideias que as cercam (Altman, 2012, p. 23). Entretanto, a explicitação e consideração desse princípio é fundamental a uma pesquisa que busca compreender o desenvolvimento de modelos sintáticos na relação incontornável com o pensamento gramaticográfico do tempo em que eles surgem.

O segundo e terceiro princípios – *imanência* e *adequação* – orientam a análise, reconstrução e sistematização do modelo sintático propriamente dito. De um lado, a *imanência* corresponde ao exame da sintaxe de Barros (1540) dentro de seu próprio quadro teórico, terminológico e conceitual, evitando anacronismos decorrentes da consideração de modelos sintáticos ou metalinguagens contemporâneas. Isso resulta, notadamente, na representação do modelo a partir de uma “rede taxonômica”, em que metatermos, relações e exemplos são tratados internamente, de maneira endógena (ver Figura 1). De outro lado, a *adequação* permite e requer aproximações entre o saber sintático de Barros e ideias linguísticas contemporâneas, em especial quanto ao quadro terminológico e conceitual da sintaxe. Esse procedimento, de acordo com Koerner (2014b, p. 89), facilita a compreensão dos resultados da

análise historiográfica por parte de linguistas, gramáticos, filólogos, professores de língua e outros leitores de hoje. No âmbito desta pesquisa, o “glossário de metatermos” é o produto central derivado da aplicação desse princípio (ver Quadro 4).

A investigação, portanto, conjuga duas dimensões de análise: uma relacionada ao conteúdo (ideias linguísticas) e outra, ao contexto (envoltura). Dessa forma, a compreensão de certos aspectos do modelo sintático de Barros (1540) pode exigir, eventualmente, a exploração do clima de opinião da época, do horizonte de retrospecção, do percurso biográfico do autor, de suas redes intelectuais e de outros instrumentos linguísticos contemporâneos à obra. Não se pode perder de vista, entretanto, que este estudo é uma historiografia da gramaticografia predominantemente orientada para o conteúdo do instrumento linguístico (dada a natureza do objeto, os propósitos da investigação e os produtos resultantes das análises), em detrimento de análises predominantemente orientadas para seu contexto de produção e circulação.

Por fim, embora a Historiografia da Linguística não siga um percurso metodológico com etapas estritamente protocolares, os procedimentos gerais empregados nesta pesquisa são comuns ao campo e resumíveis nas notórias *fases de organização do trabalho historiográfico*, apresentadas em Swiggers (2012, p. 43-44; 2013, p. 44-45; 2015, p. 12-13). Assim, a *fase heurística* da pesquisa abarcou várias etapas, incluindo o levantamento dos instrumentos linguísticos do português anteriores à publicação da primeira gramática brasileira (Moraes Silva, 1806); a verificação da disponibilidade desses instrumentos para aquisição ou consulta; a pré-análise e seleção das obras que compõem a epi-historiografia apresentada no Quadro 1; a seleção, leitura e síntese de fontes secundárias relacionadas a João de Barros, sua *Grammatica* e o contexto intelectual da época; e o mapeamento da categorização, terminologia e formação de conceitos¹⁷ para o exame do modelo sintático de Barros (1540). A *fase hermenêutica* envolveu a análise crítica e a interpretação das ideias mapeadas na fase heurística. Foram identificados os metatermos da sintaxe sob análise, tanto em termos de forma quanto de conteúdo, bem como as relações que eles estabelecem entre si, com o propósito de reconstrução e sistematização do modelo sintático por meio da montagem de uma “rede taxonômica” e da elaboração de seu “glossário de metatermos”. Essas duas ferramentas direcionam o formato da exposição historiográfica, sendo, ao mesmo tempo, processos e produtos da análise demonstrada na *fase executiva* da pesquisa e compartilhada com a comunidade acadêmica neste artigo.

2.3 Princípios, procedimentos e critérios para elaboração de redes taxonômicas e glossário de metatermos

A “rede taxonômica” e o “glossário de metatermos” representam a essência da dimensão meta-historiográfica deste trabalho. A meta-historiografia consiste na reflexão sistematizada sobre a prática historiográfica, com ênfase em seus aspectos teóricos e metodológicos (Swiggers, 2009a, p. 71). Um trabalho meta-historiográfico examina, discute, propõe e define conceitos e técnicas utilizados na historiografia, incluindo questões como periodização, interpretação evolutiva e descrição de conteúdos. É fato que a discussão e organização dos fundamentos epistemológicos realizadas, por exemplo, nesta seção configuram, *lato sensu*,

¹⁷ Na esteira de Swiggers (2012), considero a *categorização, a terminologia e a formação de conceitos* parâmetros básicos para um quadro descritivo-explicativo de uma fonte historiográfica em sua imanência.

uma meta-historiografia. No entanto, é na apresentação dessas duas ferramentas (a rede e o glossário) que tarefas de natureza “meta” são delineadas, tais quais: i) a avaliação e proposição de abordagens historiográficas para o estudo da história da sintaxe (*tarefa crítica*); ii) a elaboração de ferramentas de descrição e interpretação de modelos sintáticos (*tarefa construtiva*); e iii) reflexões sobre as formas de apresentação da Historiografia da Linguística e da historiografia da gramaticografia e da terminografia (*tarefa metateórica*)¹⁸. Trata-se de tarefas complexas, em andamento e da ordem da “meta-historiografia da gramaticografia” (Swiggers, 2020, p. 143), isto é, do campo da reflexão crítica sobre a modelização, sobre as propostas e abordagens metodológicas e epistemológicas em gramaticografia e sobre o estatuto do fazer gramatical.

Assim, a montagem da rede taxonômica para representar o modelo sintático de Barros (1540), bem como a elaboração de seu glossário de metatermos demandaram a aderência a certos princípios, procedimentos e critérios analítico-conceituais, apresentados na sequência. Eles podem ser aplicados de maneira extensiva à produção de redes e glossários de modelos sintáticos de outros instrumentos linguísticos.

Em primeiro lugar, considera-se o *conteúdo* dos metatermos na perspectiva da interface *conteúdo focal* e *conteúdo contrastivo*¹⁹, conforme apresentado em Swiggers (2010b, p. 18) e desenvolvido, por exemplo, em Polachini (2018), na identificação e análise da rede conceitual do metatermo “verbo substantivo” em gramáticas brasileiras do século 19. Nesse contexto, o conteúdo focal se refere à relação bilateral entre o significante e o significado do metatermo em si, enquanto o conteúdo contrastivo emerge da rede implícita ou explícita de conteúdos dentro da qual o metatermo assume um conteúdo dinâmico. No entanto, a rede taxonômica e o glossário de metatermos da sintaxe de Barros (1540) revelam que, em última instância, o conteúdo de um metatermo é sempre relacional. Por exemplo, a compreensão do metatermo “trāſituo” [transitivo] no modelo sintático em questão exige a consideração de processos relacionais de hierarquização, classificação e contraste envolvendo metatermos como “peſoál” [pessoal], “vérbo” [verbo] e “neutro”. Como será visto adiante, processos relacionais como esses aparecem diagramados na rede da Figura 1 e explicitados no glossário do Quadro 6 (ver verbete 69).

Em segundo lugar, é importante reconhecer que uma terminografia deve lidar com metatermos em diferentes níveis de referência, desde os mais “substanciais” até os mais “relacionais”, conforme discutido em Swiggers (2009b, p. 19) e Gómez Asencio e Gonçalves (2015, p. 73). Assim, os critérios de classificação de metatermos presentes nesses estudos foram expandidos, consolidados, reorganizados e adaptados à análise de modelos sintáticos, resultando em seis tipos de metatermo, conforme apresentado no quadro seguinte:

¹⁸ A reflexão originária sobre esses três tipos de tarefa da meta-historiografia – tarefa crítica, construtiva e metateórica – pode ser conferida em Swiggers (2004, p. 116-117; 2009a, p. 71; 2010a, p. 5; 2013, p. 40).

¹⁹ O *conteúdo* dos metatermos é apenas um dos sete critérios de análise de metatermos apresentados em Swiggers (2010b, p. 18-19). Os demais critérios são a *incidência* dos metatermos e sua *abrangência heurística, teórica, disciplinar, macrocientífica e cultural*. Quando oportuno, esses valores, denominados pelo autor de “parâmetros classêmicos”, também foram considerados na análise do modelo sintático de Barros (1540).

Quadro 4 – Níveis de referência dos metatermos em modelos sintáticos

Tipos	Referentes	Exemplos (Barros, 1540)
Unidades	Classes e subclases de palavras, categorias e subcategorias sintáticas, constituídas por conjuntos paradigmáticos de elementos concretos.	ajetiuo [adjetivo], prepoſiçam [preposição], v̄rbo tr̄ſituo [verbo transitivo]
Acidentes	Traços ou propriedades que se sobrepõem ou afetam as unidades.	caſo [caso], numero [número], peſoa [pessoa]
Relações	Ligações formais ou lógicas entre unidades (ou acidentes), funções sintáticas, processos e procedimentos gramaticais.	appoſitio [aposição], concordânciā [concordância], regimento
Construções	Composições entre unidades, constituindo outras unidades ou fenômenos sintáticos.	polyſyntheton [polissíndeto], uóz paſſiua [voz passiva], zeuma [zeugma]
Dêixis	Elementos que fazem referência ao próprio modelo sintático, a seus objetos e a suas partes principais.	conſtruiçám [construção], figura, syntaxis [sintaxe]
Outros	Resíduos da descrição/prescrição sintática, elementos não classificados, vocabulário comum (não técnico).	v̄rbo ey [verbo hei], pronomē se

Fonte: Elaboração própria.

Um terceiro ponto a ser considerado na análise de um modelo sintático é a natureza nem sempre transparente da relação entre metatermos e conceitos num instrumento linguístico. Como antecipado em Swiggers (2009b, p. 20), a relação entre o metatermo e o conceito pode ser confusa, ou o próprio conceito pode ser pouco preciso. Na análise terminográfica da sintaxe de Barros (1540), são exemplos dessa opacidade os metatermos “scheſionomaton” e “dyaleton”, que designam tipos de solecismo. No glossário de metatermos, essas entradas não foram apresentadas ao lado de metatermos contemporâneos correspondentes ou simplesmente com as grafias atualizadas, haja vista: i) o desconhecimento, por parte deste historiógrafo, de um metatermo em português contemporâneo equivalente ao metatermo “scheſionomaton”, ou mesmo de uma adaptação ortográfica em uso; ii) a distância conceitual entre o metatermo “dyaleton”, presente em Barros (1540), e o metatermo contemporâneo “dialeto”. Além disso, a relação “metatermo-conceito” nem sempre é unívoca (de um para um), tornando-se por vezes equívoca: um metatermo pode corresponder a vários conceitos, como é o caso de “conſtruiçám” [construção], que designa tanto a sintaxe quanto seu objeto (ver verbete 20, Quadro 6); ou um conceito pode corresponder a vários metatermos, como nos casos dos pares “comum” e “apellatiuo” [apelativo], ou “figura” e “uiçō” [vídeo].

Observações dessa natureza demandaram a aplicação de alguns procedimentos e critérios no percurso da elaboração do glossário de metatermos de Barros (1540), a fim de evitar distorções de significados nos verbetes e incongruências entre a rede taxonômica e o glossário. Desse modo, a abordagem inicial procurou entender o conceito dos metatermos em seu contexto original de produção, circulação e recepção, antes de estabelecer comparações com metatermos e conceitos contemporâneos. Em alguns casos, formas e construções arcaicas, características da elocução gramaticográfica renascentista, foram modificadas para uma linguagem atual, tentando preservar, não obstante, a integridade e historicidade dos metatermos. Por essa razão, os verbetes não são elaborados apenas a partir de definições ou conceitos

retirados diretamente do instrumento linguístico²⁰, mas principalmente por meio da relação em rede com os outros verbetes do modelo sintático (conteúdo contrastivo imanente) e de comentários direcionados ao leitor do presente (princípio da adequação). Trata-se, portanto, de um glossário mais explicativo que meramente definicional, que mescla a voz do gramático renascentista à de seu historiógrafo, proporcionando ao pesquisador ou consultante contemporâneo uma visão mais elucidativa dos metatermos.

Em quarto lugar, gostaria de apresentar, de forma concisa, alguns procedimentos finais, igualmente planejados e executados durante a análise do modelo sintático de Barros (1540), bem como na criação da rede taxonômica e do glossário de metatermos. É importante destacar que esses procedimentos podem ser replicados na análise e sistematização de modelos sintáticos de outros instrumentos linguísticos. Ei-los:

- 1 Quando ocorre uma contradição entre uma afirmação retórica e uma descrição/normatização gramatical (com ou sem exemplos), a prioridade é dada à descrição/normatização gramatical na reconstrução do modelo sintático. Ou seja, são considerados os casos efetivamente gramatizados pelo instrumento linguístico.
- 2 São ignorados os metatermos que, apesar de mencionados na obra (seja para rejeitá-los ou para descrever/normatizar outros níveis gramaticais, por exemplo), não fazem parte das análises metalingüísticas ou das regras normativas que compõem o conhecimento sintático presente no material.
- 3 Por outro lado, são incluídos os metatermos que, embora não apareçam nas partes, capítulos ou seções explicitamente dedicadas à sintaxe, são mencionados em outras partes, capítulos ou seções e fazem parte indubitavelmente do conjunto de saberes sintáticos da obra, no eixo da análise metalingüística e/ou da norma-padrão.
- 4 Os metatermos da rede taxonômica são apresentados em seu sistema gráfico original, incluindo o uso de letras como “ß” [s germânico ou *eszett*] e “ſ” [s longo], e de combinações de letras e diacríticos como “é” [e com cedilha] e “ë” [e com til], as quais não fazem parte do sistema gráfico atual da língua portuguesa. Da mesma forma, a grafia original das entradas dos verbetes do glossário é mantida, e, quando necessário, esses metatermos são acompanhados por suas respectivas grafias atualizadas, apresentadas entre colchetes.
- 5 São preservadas também a grafia (incluindo o uso de maiúsculas) e a pontuação original dos exemplos na rede taxonômica. Esses exemplos são transcritos de forma integral ou parcial, dependendo da natureza do que se pretende exemplificar (uma regra de concordância, o caso vocativo, um tipo de solecismo etc.).
- 6 A rede taxonômica é diagramada utilizando o site “canva.com” e organizada com a ajuda de símbolos que atribuem significados específicos ao modelo sintático representado. Os símbolos e seus respectivos significados estão reunidos no Quadro 5:

²⁰ “A menos que o único objetivo do historiógrafo seja colecionar antiguidades, isto é, descrever conceitos desenvolvidos muitos anos atrás unicamente nos próprios termos utilizados, ele será tentado a usar um vocabulário técnico moderno na sua análise” (Koerner, 1996b, p. 98).

Quadro 5 – Legenda para leitura de redes taxonômicas de representação de modelos sintáticos

Símbolo	Descrição do símbolo	Significado
LOREM (2023)	Retângulo preto	Autor e ano do instrumento linguístico
lorem ipsum	Retângulo cinza escuro	Parte da macroorganização do modelo sintático
lorem ipsum	Retângulo cinza claro	Unidade primária (ou única) de uma construção
lorem ipsum	Retângulo branco	Unidade (ou acidente) secundária de uma construção
lorem ipsum	Retângulo tracejado	Facultatividade ou suposição
lorem ipsum	Retângulo cinza escuro com borda arredondada	Objeto/unidade de análise
lorem ipsum	Retângulo cinza claro com borda arredondada	Particularidade descritiva/normativa
lorem ipsum	Retângulo branco com borda arredondada	Acidente
lorem ipsum	Retângulo tracejado com borda arredondada	Exemplo
↔	Seta bilateral	Relação (formal ou lógica) entre unidades (ou entre unidade e acidente)
↔→	Seta bilateral tracejada	Relação facultativa (formal ou lógica) entre unidades (ou entre unidade e acidente)
○—○	Linha com círculos nas pontas	Relação de subclassificação/subcategorização
-----	Linha tracejada	Equivalência
1	Círculo branco numerado	Enumeração de tipos
●	Círculo azul	Unidade, acidente ou construção semelhante à ideia de “sujeito”
○	Círculo amarelo	Unidade, acidente ou construção semelhante à ideia de “verbo”
●	Círculo verde	Unidade, acidente ou construção semelhante à ideia de “complemento verbal”
●	Círculo rosa	Unidade, acidente ou construção semelhante à ideia de “adjunto adverbial”
●	Círculo verde-limão	Exemplo semelhante à ideia de “orações coordenadas”
●	Círculo vermelho	Exemplo semelhante à ideia de “oração subordinada substantiva”
●	Círculo laranja	Exemplo semelhante à ideia de “oração subordinada adjetiva”
●	Círculo roxo	Exemplo semelhante à ideia de “oração subordinada adverbial”

Fonte: Elaboração própria.

Por fim, devo fazer uma observação complementar à legenda apresentada no Quadro 5. A ordem de tipos na rede taxonômica, simbolizada por círculos brancos numerados, nem sempre corresponde à disposição desses tipos no instrumento linguístico. Tenta-se seguir a ordem apresentada na obra, mas essa ordem também pode ser ajustada para melhorar a organização e diagramação do modelo sintático, ou com o intuito de destacar equivalências – como exemplo, observe-se na Figura 1 a linha tracejada que conecta os metatermos “verbo” das regras (1) de concordância e regimento.

3 O primeiro modelo sintático da gramaticografia portuguesa: Barros (1540)

A codificação e normatização dos vernáculos europeus, fundamentadas na tradição greco-latina, representaram um processo gradual que teve início na Idade Média e ganhou considerável impulso a partir do Renascimento. Nesse período, essas línguas começaram a substituir progressivamente o latim nas esferas literária, administrativa e científica. Esse movimento singular na história da linguística se relaciona ao fortalecimento dos reinos europeus e à formação das línguas de cultura escrita específicas de cada um deles, envolvendo a produção de instrumentos linguísticos como gramáticas, dicionários, ortografias e manuais de tradução. Assim, a gramática deixou de ser exclusivamente latina, passando também a contemplar os vernáculos. Esse fluxo gramaticográfico massivo é abordado detalhadamente em Auroux (2014), que o denomina “revolução tecnológica da gramatização”.

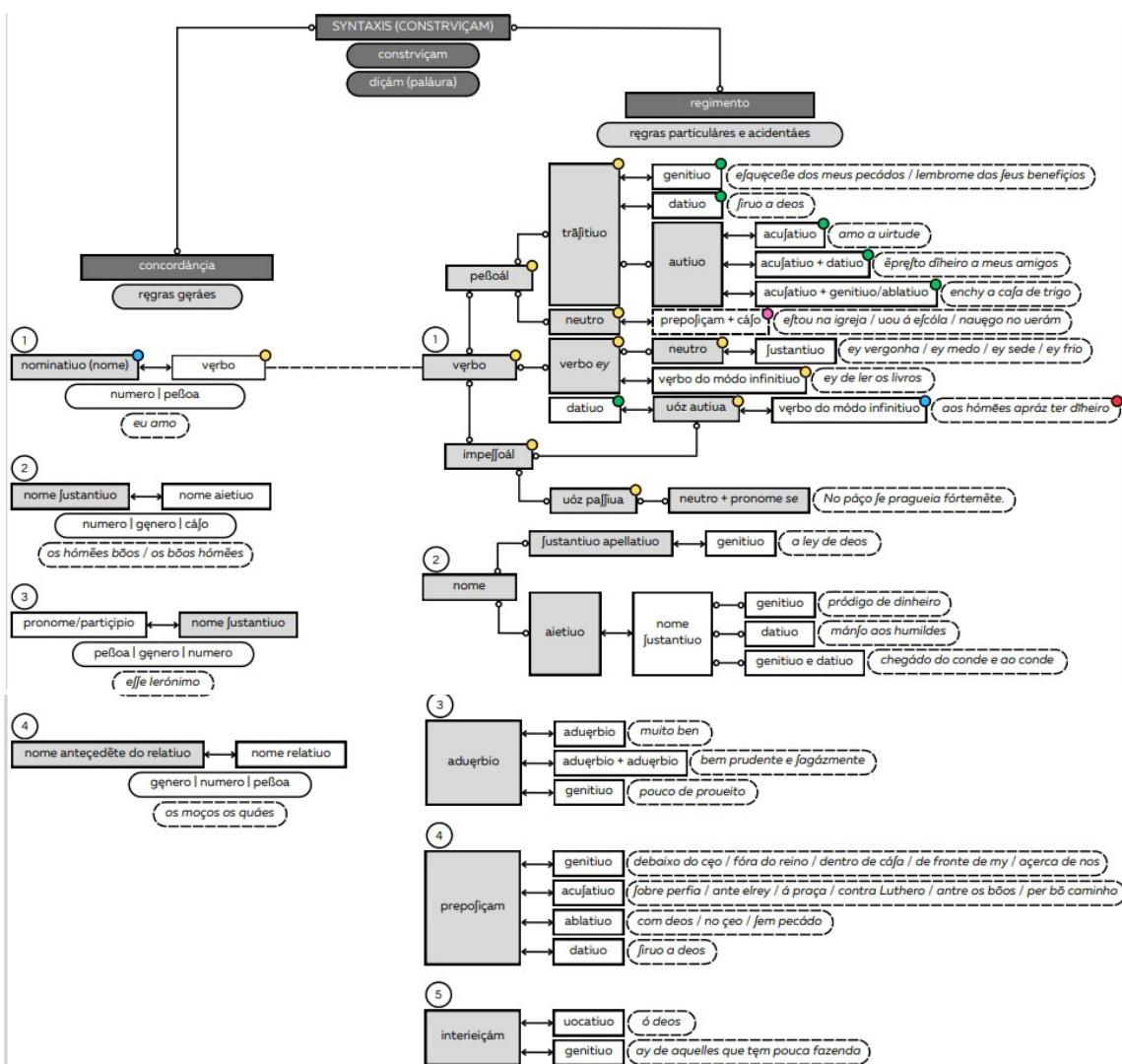
No contexto português, João de Barros, com sua pioneira sistematização descriptivo-normativa, desempenhou papel basilar nessa empreitada gramaticográfica. Nas primeiras décadas do século 16, a língua portuguesa já estava estabelecida como língua escrita de um império em expansão, mas ainda precisava se afirmar diante do latim e, cada vez mais, em relação ao espanhol, que desempenhava um papel significativo como língua de cultura em Portugal (Schäfer-Prieß, 2019, p. 84). A obra de Barros serviu para afirmar a autonomia do português no período dos quinhentos, documentando a integração da atividade intelectual portuguesa na cultura do Ocidente europeu (Buescu, 1984, p. 63).

O humanista, político, historiador e gramático João de Barros nasceu por volta de 1496, em Vila Verde ou Viseu, Portugal. De família fidalga e órfão desde tenra idade, ele foi acolhido nos Paços da Ribeira, o palácio real, onde desempenhou a função de “moço de guarda-roupa” para o futuro rei D. João III (1502-1557). Sua educação na corte o tornou um homem culto, conhecedor das línguas clássicas e suas literaturas. Ao longo da vida, ocupou vários cargos públicos em destaque na política portuguesa e, simultaneamente, dedicou-se à atividade de escritor. Em 20 de dezembro de 1539, publicou seu primeiro instrumento linguístico, a *Grammatica da lingua portuguesa com os mandamentos da Santa madre igreja*, uma cartilha contendo preceitos, mandamentos e um missal destinados ao ensino das primeiras letras. Esse material serviria como introdução à *Grammatica da lingua Portuguesa*, publicada dias depois (em 1540) e acrescida de dois diálogos em estilo socrático: *Dialogo em Lovvor da nossa Lingvagem* e *Dialogo da Viçosa Vergonha*, com fundamentos apologéticos e morais, respectivamente. Como os títulos sugerem, o conjunto de livros destinava-se à formação geral e grammatical dos aprendizes, doutrinando-os na fé cristã e nos bons costumes. Posteriormente, Barros ainda rascunhou uma gramática latina intitulada *Grammatices Rudimenta ou Humilia*

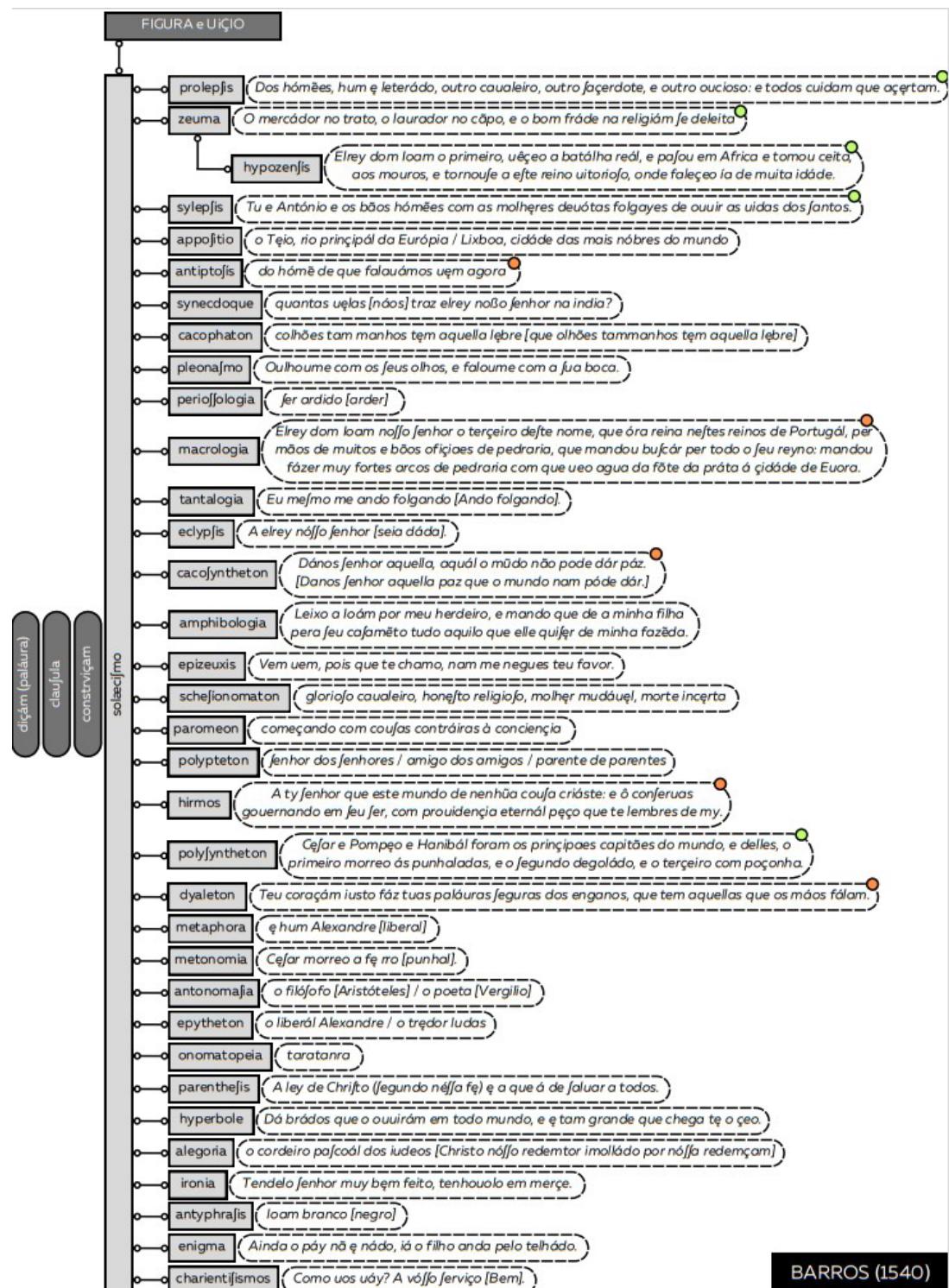
Praecepta, mas não chegou a concluir. Em 1568, o gramático se afastou de suas funções oficiais e se retirou para Pombal, em Portugal, onde veio a falecer em 1570 ou 1571.²¹

Dentre esses instrumentos linguísticos, é na *Grammatica da lingua Portuguesa* (Barros, 1540) que se encontra o primeiro modelo sintático da gramaticografia do português. A seguir, a Figura 1 e o Quadro 6 analisam, reconstruem e sistematizam esse modelo. A fonte do trabalho é a primeira edição da obra, publicada em 1540, em Lisboa. Outras três edições subsequentes, publicadas em 1785, 1957 e 1971, também em Lisboa, não apresentaram alterações no conhecimento sintático contido no instrumento:

Figura 1 – Rede taxonômica do modelo sintático de Barros (1540)



²¹ Informações biográficas coletadas em Buescu (1984, p. 29-32), Casagrande (2005, p. 71-73), Fernandes (2005, p. 131), Leite (2007, p. 87-91) e Schäfer-Prieß (2019, p. 12-13).



Fonte: Elaboração própria.

Quadro 6 – Glossário dos metatermos do modelo sintático de Barros (1540)

1. **ablatiuo** [ablativo]: sexto caso dos nomes (os demais são nominativo, genitivo, dativo, acusativo e vocativo), usado quando se tira ou se afasta a coisa de algum lugar.
2. **acusatiuo** [acusativo]: quarto caso dos nomes (os demais são nominativo, genitivo, dativo, vocativo e ablativo), no qual se põe a coisa feita ou amada.
3. **aduérbio** [advérbio]: parte da oração (as demais são artigo, nome, pronome, verbo, particípio, conjunção, preposição e interjeição) que sempre anda junta e costurada ao verbo, de onde vem seu nome: “ad” significa “cerca” e “advérbio”, “acerca do verbo”; denota a eficácia ou remissão do verbo, pois tem o poder de acrescentar, diminuir ou destruir totalmente a ação do verbo ao qual se junta, lhe dando quantidade ou qualidade acidental, como o adjetivo ao substantivo.
4. **aietiuo** [adjetivo]: qualidade do nome que se distingue da qualidade “substantivo” por não ter existência em si, mas estar ligado a este e poder ser antecedido pela palavra “coisa”.
5. **alegoria**: espécie de solecismo que quer dizer “significação alheia”; há seis espécies de alegoria, e a primeira ocorre quando entendemos uma coisa por meio de outra.
6. **amphibologia** [anfibologia]: espécie de solecismo que quer dizer “dúvida de palavras”, por meio das quais duvidamos da sentença em que elas estão, o que resulta muitas vezes em grandes controvérsias e disputas.
7. **anteçedête** [antecedente]: qualidade do nome que se distingue da qualidade “relativo” por ser o nome que fica atrás deste e que por ele é lembrado.
8. **antiptofis** [antiptose]: espécie de solecismo que quer dizer “caso por caso”; ocorre quando colocamos em um determinado caso uma coisa que deveria estar em outro.
9. **antonomafia** [antonomásia]: espécie de solecismo que quer dizer “substituição de um nome por outro”; ocorre quando usamos um nome comum em lugar de um nome próprio, devido a alguma excelência deste.
10. **antyphrafis** [antífrase]: espécie de solecismo que quer dizer “fala contrária”; ocorre quando entendemos um nome por meio de outro contrário a ele.
11. **apellatiuo** [apelativo]: qualidade do nome pela qual entendemos muitas coisas de um determinado gênero; também denominada “comum”, opondo-se a “próprio”.
12. **appositio** [aposição]: espécie de solecismo que quer dizer “justaposição”; ocorre quando juntamos dois nomes substantivos sem conjunção, em que um expõe e declara o outro.
13. **autiuo** [ativo]: gênero do verbo pessoal (o outro é o gênero neutro) que podemos converter para o modo passivo e por meio do qual denotamos alguma ação que passe a outra coisa, posta no caso acusativo.
14. **cacophaton** [cacófato]: espécie de solecismo que quer dizer “mal som”; vício que a orelha recebe mal; comete-se quando do fim de uma palavra ao princípio de outra há alguma fealdade ou torpeza.
15. **cacosyntheton** [cacossínteto]: espécie de solecismo que quer dizer “má composição”; ocorre quando, pretendendo-se elegância, ordena-se a linguagem de acordo com o latim.
16. **cáso** [caso]: termo por onde os nomes, pronomes e participios podem andar e que, desde que não mude a substância do nome, governa a ordem da oração mediante o verbo.
17. **charientísmos** [carentismo]: espécie de solecismo que quer dizer “graciosidade”; equivale a responder com graça e benevolência, como quando nos perguntam “como vós estais?” e respondemos “ao vosso serviço”, em vez de “bem”.
18. **claufula** [cláusula]: palavras dispostas entre dois pontos; se encerra por meio de um ponto e pode ser cortada em duas partes por uma vírgula; divide com a palavra e a construção o papel de unidades de análise das figuras; também está presente nas regras de pontuação.
19. **concordância** [concordância]: uma das duas divisões da construção (a outra é o regimento); conveniência de duas dições que se correspondem em número, gênero, caso e/ou pessoa, de acordo com regras gerais.
20. **constrviçam** [construção]: quarta parte da gramática (as três primeiras são letra, sílaba e dição); é uma conveniência entre partes, postas em seus lugares naturais, por meio da qual obtemos o conhecimento dos nossos conceitos; divide-se em concordância (regras gerais) e regimento (regras particulares e acidentais); é tratada pela sintaxe, mas também corresponde a esta; divide com a palavra e a cláusula o papel de unidades de análise das figuras; outras grafias: “cõstruçã”, “côstruiçã”, “conjtruiçã”, “constrviçam”, “côstruiçã”, “conjtruiçãm”, “construiçãm”, “construiçam”.
21. **datiuo** [dativo]: terceiro caso dos nomes (os demais são nominativo, genitivo, acusativo, vocativo e ablativo), no qual pombos a pessoa em cujo proveito ou dano é dada ou feita a coisa.

22. **diçám** [dição]: objeto de estudo da Etimologia, composto por letras e sílabas; divide com a cláusula e a construção o papel de unidades de análise das figuras; o mesmo que “palavra”; outra grafia: “diçam”.
23. **dyaleton**: espécie de solecismo que quer dizer “dissolução” ou “desatamento”; ocorre quando muitas partes e cláusulas se juntam sem conjunção.
24. **eclyp̄sis** [elipse]: espécie de solecismo que quer dizer “desfalecimento”; trata-se de uma figura muito comum, principalmente nos sobreescritos das cartas.
25. **enigma**: espécie de solecismo que quer dizer “pergunta escura”; usamos quando dizemos alguma coisa por meio de palavras escuras, de modo semelhante aos jogos de adivinhação dos meninos.
26. **epizeuxis** [epizéuxis]: espécie de solecismo que quer dizer “conjunção”; ocorre quando se repete algo duas ou três vezes sem entreposição de parte.
27. **epytheton** [epíteto]: espécie de solecismo que quer dizer “colocação abaixo do nome”; cometemos essa figura quando, por meio de um nome adjetivo, queremos elogiar ou depreciar alguém ou algo.
28. **figura**: forma de expressão por meio de alguma arte inovadora; é também denominada “vício”; as figuras se dividem em dois gêneros – barbarismo e solecismo – e tais gêneros, em várias espécies.
29. **gênero** [gênero]: acidente do nome (os demais são qualidade, espécie, figura, número e declinação por caso), do pronome (os demais são espécie, número, figura, pessoa e declinação por caso) e do verbo (os demais são espécie, figura, termo, modo, pessoa, número e conjugação); o gênero do nome é uma distinção pela qual reconhecemos o macho da fêmea, e o neutro de ambos; o gênero do pronome pode ser masculino, feminino, neutro e comum de dois; o gênero do verbo, que pode ser ativo ou neutro, é uma natureza especial que alguns verbos têm (os pessoais) e outros não; as figuras também são divididas em gêneros: barbarismo e solecismo.
30. **genitíuo** [genitivo]: segundo caso dos nomes (os demais são nominativo, dativo, acusativo, vocativo e ablativo), assim denominado por gerar os outros casos latinos; é também denominado “caso possessivo e interrogativo”, por nele estar o senhor da coisa: se perguntassemos “de quem é esta arte de gramática?”, a resposta seria “do Príncipe, nosso Senhor”.
31. **hirmos**: espécie de solecismo que quer dizer “estendimento”; cometemos essa figura quando prolongamos uma sentença com grande arrazoamento de palavras, e em seguida arrematamos.
32. **hyperbole** [hipérbole]: espécie de solecismo que quer dizer “excedimento”; comete-se quando, para elogiar ou depreciar alguma coisa, se diz algo que ultrapassa a verdade.
33. **hypozēfis**: espécie de solecismo que quer dizer “ajuntamento debaixo”; espécie de zeugma e contrária a ela.
34. **impejoál** [impessoal]: verbo que só se conjuga na terceira pessoa do número singular e não tem primeira nem segunda pessoa, ao contrário do verbo pessoal; se apresenta na voz ativa ou passiva.
35. **infinitíuo** [infinitivo]: quinto modo do verbo (os outros são indicativo, imperativo, optativo e subjuntivo), que quer dizer “não acabado”, porque, além de carecer de número e pessoa, não determina nem acaba coisa alguma sozinho; nos permite conhecer a conjugação de qualquer verbo e tomar as regras para formação dos outros; quando antecedido por artigo, pode ser considerado “nome verbal”; também denominado “infinito”, se opõe ao modo finito.
36. **interieicâm** [interjeição]: parte da oração (as demais são artigo, nome, pronome, verbo, advérbio, particípio, conjunção e preposição) que denota o que a alma padece; rege o caso vocativo; outra grafia: “interieicã”.
37. **ironia**: espécie de solecismo que quer dizer “dissimulação”; usamos quando, por meio do contrário, dizemos o que queremos; nesse sentido, os gestos do corpo e o tom da voz auxiliam nossa intenção.
38. **macrologia**: espécie de solecismo que quer dizer “longo rodeio de palavras”; comete-se quando contamos algo e damos muitas voltas para concluir uma sentença.
39. **metaphora** [metáfora]: espécie de solecismo que quer dizer “transformação”; usamos quando atribuímos a uma coisa alguma conveniência ou especialidade de outra coisa.
40. **metonomia** [metonímia]: espécie de solecismo que quer dizer “transnominación”; comete-se quando substituímos o instrumento pela coisa com que ele é feito, ou a matéria por aquilo que se faz dela.
41. **módo** [modo]: acidente do verbo pessoal (os demais são gênero, espécie, figura, tempo, pessoa, número e conjugação) que denota a vontade ao falar; são cinco os modos do verbo: indicativo, imperativo, optativo, subjuntivo e infinitivo.
42. **neutro**: gênero do verbo pessoal (o outro é o gênero ativo) que não se pode converter para o modo passivo (ao contrário do verbo ativo) e cuja ação não passa a outra coisa (ao contrário do verbo transitivo); não requer caso depois de si, exceto mediante preposição; também pode se referir a um gênero do nome e do pronome, distinguindo-se dos gêneros masculino, feminino e comum de dois.

43. **nome**: parte da oração (as demais são artigo, pronome, verbo, advérbio, participípio, conjunção, preposição e interjeição) que se declina por casos, não tem tempo e significa sempre alguma coisa com ou sem corpo; tem os seguintes acidentes: qualidade, espécie, figura, gênero, número e declinação por caso.
44. **nominatiuo** [nominativo]: primeiro caso dos nomes (os demais são genitivo, dativo, acusativo, vocativo e ablativo), por ser o primeiro que nomeia a coisa que é ou a pessoa que faz.
45. **numero** [número]: acidente do nome (os demais são qualidade, espécie, figura, gênero e declinação por caso), do pronome (os demais são espécie, gênero, figura, pessoa e declinação por caso) e do verbo pessoal (os demais são gênero, espécie, figura, tempo, modo, pessoa e conjugação), que pode ser singular ou plural; número do nome é a distinção entre um e muitos.
46. **onomatopeia**: espécie de solecismo que quer dizer “fingimento de nome”.
47. **paláura** [palavra]: o mesmo que “diçam”; outra grafia: “palaúra”.
48. **parenthesis** [parêntese]: espécie de solecismo que quer dizer “entreposição”; usa-se quando, no meio de alguma sentença, se entrepõem outras palavras fora do seu propósito.
49. **paromeon**: espécie de solecismo que quer dizer “semelhante princípio”; comete-se essa figura quando se começam muitas dições por uma mesma letra.
50. **particiípio** [participípio]: parte da oração (as demais são artigo, nome, pronome, verbo, advérbio, conjunção, preposição e interjeição); discurso por meio do qual o verbo na voz ativa é conjugado (os demais são o discurso presente, pretérito, infinitivo e gerúndio); é também denominado “participípio passado” e “participípio do pretérito”.
51. **perioffologia** [perissologia]: espécie de solecismo que quer dizer “sobejidão de razões”; cometemos essa figura quando, por meio de palavras dobradas que não têm mais força, dizemos algo que pode ser dito com poucas palavras.
52. **peñoa** [pessoa]: acidente do pronome (espécie, gênero, número, figura e declinação por caso) e do verbo pessoal (os demais são gênero, espécie, figura, tempo, modo, número e conjugação) classificado em três: primeira, segunda e terceira.
53. **peñoál** [pessoal]: verbo que tem número e pessoa, ao contrário do verbo impessoal; traz consigo oito acidentes: gênero, espécie, figura, tempo, modo, pessoa, número e conjugação.
54. **pleonasmo** [pleonasmo]: espécie de solecismo que quer dizer “sobejidão de palavras”; cometemos essa figura quando dizemos algumas palavras que poderiam ser dispensadas.
55. **polypteton** [poliptoto]: espécie de solecismo que quer dizer “multiplicidade de casos”; ocorre quando juntamos casos distintos.
56. **polysyntheton** [polissíndeto]: espécie de solecismo que quer dizer “composição de muitos”; comete-se quando muitas palavras e cláusulas se juntam por conjunção.
57. **preposiçam** [preposição]: parte da oração (as demais são artigo, nome, pronome, verbo, advérbio, participípio, conjunção e interjeição) que se põe entre as outras partes por ajuntamento ou composição. Pode reger o caso genitivo, acusativo, ablativo ou dativo; outra grafia: “prepojiçã”.
58. **prolepsis** [prolepsis]: espécie de solecismo que quer dizer “antecipação”. Comete-se quando se divide alguma generalidade em diversas partes.
59. **pronomé**: parte da oração (as demais são artigo, nome, verbo, advérbio, participípio, conjunção, preposição e interjeição) que se põe no lugar do próprio nome; tem seis acidentes: espécie, gênero, número, figura, pessoa e declinação por caso.
60. **regimento** [regência]: uma das duas divisões da construção (a outra é a concordância); construção de duas dições diferentes em gênero, número, caso ou pessoa, em que, apenas por uma natureza especial, uma obriga a seguir a serposta num determinado caso, de acordo com regras particulares e acidentais.
61. **relatiuo** [relativo]: qualidade do nome que se distingue da qualidade “antecedente” por ser a parte que lembra algum nome que fica atrás, isto é, a parte que lembra seu antecedente.
62. **scheinonomaton**: espécie de solecismo que quer dizer “confusão de nomes”, como quando, para encher a oração, juntamos muitos substantivos e adjetivos.
63. **solecismo** [solecismo]: segundo gênero das figuras ou vícios (o primeiro é o barbarismo), que podemos cometer na construção e na ordem das partes, quando as usamos de alguma maneira diferente da fala comum.
64. **sylepjis** [silepse]: espécie de solecismo que quer dizer “concebimento”; ocorre quando, por meio de nomes substantivos e adjetivos de números diferentes, e pronomes de pessoas diferentes, colhemos com um verbo uma cláusula.

65. **synecdoque** [sinédoque]: espécie de solecismo que quer dizer “entendimento”; ocorre quando entendemos o todo pela parte.
66. **syntaxis** [sintaxe]: uma das quatro partes da gramática (as outras são ortografia, prosódia e etimologia); tanto corresponde à construção quanto trata dela.
67. **fustantiuo** [substantivo]: qualidade do nome que se distingue da qualidade “adjetivo” por ter existência em si e não poder ser antecedido pela palavra “coisa”.
68. **tantalogia** [tautologia]: espécie de solecismo que quer dizer “repetição de uma palavra muitas vezes”.
69. **trāsituo** [transitivo]: verbo pessoal que quer dizer “passador”, ou seja, que passa sua ação a outra coisa (ao contrário do verbo neutro); pode reger diferentes casos.
70. **uocatiuo** [vocativo]: quinto caso dos nomes (os demais são nominativo, genitivo, dativo, acusativo e ablativo), no qual está a pessoa que chamamos; é regido pela interjeição “ó” e outras que se veem em seu lugar; outra grafia: “vocatiuo”.
71. **uiçio** [vício]: o mesmo que “figura”.
72. **verbo** [verbo]: parte da oração (as demais são artigo, nome, pronome, advérbio, particípio, conjunção, preposição e interjeição) que não se declina por casos como o nome e o pronome, mas se conjuga em modos e tempos; voz ou palavra que demonstra fazer alguma coisa; divide-se, por um lado, em substantivo (verbo *ser*) e adjetivo (demais verbos) e, por outro, em pessoal e impersonal.
73. **vóz autiua** [voz ativa]: voz ou modo que age ao realizar a ação; opõe-se à voz passiva e, assim como ela, é também uma das duas maneiras de ser do verbo impersonal.
74. **vóz paſſiua** [voz passiva]: voz ou modo que padece ao receber a ação; modo em que o verbo ativo pode ser convertido (ao contrário do verbo neutro); supre, por rodeio, com o verbo *ser* e um particípio do tempo passado, a ausência de verbo passivo em português; opõe-se à voz ativa e, assim como ela, é também uma das duas maneiras de ser do verbo impersonal, sendo formado, nesse caso, por um verbo neutro e o pronome de terceira pessoa *se*.
75. **zeuma** [zeugma]: espécie de solecismo que quer dizer “conjunção”. Ocorre quando damos muitas partes a um verbo. Contrário de “*hypozenſis*”.

Fonte: Elaboração própria.

A rede taxonômica (Figura 1) é uma espécie de fotografia da sintaxe presente na *Grammatica da lingua Portuguesa*, de João Barros (1540). O glossário de metatermos que a acompanha (Quadro 6) pode ser consultado para esclarecer dúvidas durante a leitura da rede ou para compreender o significado de um metatermo num sistema terminológico particular e na relação com o contexto intelectual em que o instrumento linguístico se insere. Apesar de ambas as ferramentas serem, em grande parte, autoexplicativas, alguns aspectos gerais e específicos serão retomados e desenvolvidos na sequência.

3.1 Inspiração em Nebrija e macro-organização sintática

Costuma-se afirmar que a *Grammatica de la lengua castellana*, publicada em 1492 pelo filólogo espanhol Elio Antonio de Nebrija (1444-1522), é a principal fonte de inspiração para a *Grammatica* de João de Barros (cf. Buescu, 1984, p. 87; Leite, 2007, p. 105). Muito por isso, o autor português é frequentemente considerado pouco original pela literatura especializada, sobretudo quando comparado a Fernão d’Oliveira (Schäfer-Prieß, 2019, p. 14). É de se esperar, portanto, que a macro-organização do modelo sintático de Barros reflita o que há no Livro IV (“Que es de sintaxis y orden de las diez partes de la oración”) da gramática de Nebrija: inicialmente, a exposição das relações e regras de concordância e regência (Capítulos I a IV do Livro IV); em seguida, a apresentação de uma sintaxe irregular, centrada em uma lista de solecismos (Capítulos V e VII do Livro IV).

De fato, é essa estrutura que se vê na Figura 1, em que a sintaxe de Barros (1540) é apresentada em duas partes, isto é, em dois grandes grupos de metatermos em rede. O primeiro grupo corresponde à sintaxe propriamente dita, que se subdivide em regras gerais de concordância (quatro tipos) e regras particulares e accidentais de regimento [regência] (cinco tipos), tratadas principalmente no capítulo “Da construicām das pártes” (Barros, 1540, p. 30-33 verso). O segundo grupo abrange um conjunto de 34 tipos de solecismo²², arrolados no capítulo “Das figvras” (Barros, 1540, p. 34-39 verso), na mesma sequência em que aparecem na obra de Nebrija (1492, p. 50-54)²³. Essa macro-organização, a propósito, persistiu na maioria das obras gramaticais em português subsequentes, pelo menos até o final do século 18, evidenciando a permanência das características centrais do modelo sintático de Barros na gramaticografia da língua.

Portanto, João de Barros incorporou com sucesso, indiscutivelmente, elementos da sintaxe de Antonio de Nebrija em sua obra. Não se deve esquecer, no entanto, que essas semelhanças também decorrem de um patrimônio comum de gramatização de base latina, àquilo que Auroux (2014, p. 46) chama de “gramática latina estendida”. Em Barros (1540), isso pode ser observado, por exemplo, nas referências genéricas aos “latinos” e “gramáticos”, bem como a autores latinos específicos, como Quintiliano (35-96 EC), sugerindo que o autor consultou diretamente fontes latinas, como ilustram os seguintes trechos retirados da parte da gramática dedicada à sintaxe:

[...] fica agóra uermos a quārta que é da cōstruiçā, Esta (segundo difinçā dos grāmáticos) é hūa cōueniēcia antre pártes, pótas é ſeus naturáes lugáres (p. 30); Nós tomaremos da nōſſa cōstruicām o mais neceſário, immitando ſempre a órdem dos latinos (p. 30); Estes uerbos peſãoes, ou páſa a ſua auçām em outra couſa, ou nam. Os que paſſam chamālhe os latinos trāſitios (verso p. 31); Figura (ſegundo difinçā de Quintiliano) é hūa fórmā de dizer per algūa árte noua (p. 34); Muitas outras figuras tem os latinos as quáes nam exemplificamos em nōſſa lingüágem (verso p. 39, *grifos nossos*).

Além disso, de acordo com Buescu (1984, p. 67), em geral, a conformidade com o esquema dos gramáticos latinos na *Gramatica de la lengua castellana* é mais nítida e rígida do que na *Grammatica da lingua Portuguesa*, cuja latinização é mais formal do que essencial: uma análise estatística do texto de Barros revela que ele se preocupa mais em demonstrar diferenças (eles latinos vs. nós portugueses) do que apontar identidades. A ordenação da sintaxe e a organização de seus capítulos também indicam diferenças entre Barros e Nebrija, o que reforça a tese de que ele deve ter consultado fontes diferentes, além da obra castelhana. A *Grammatica* de Barros é, assim, uma resposta à iniciativa gramaticográfica de Nebrija (e não uma mera transposição de seu projeto), documentando a integração da atividade intelectual portuguesa na vida e na cultura do Ocidente europeu.

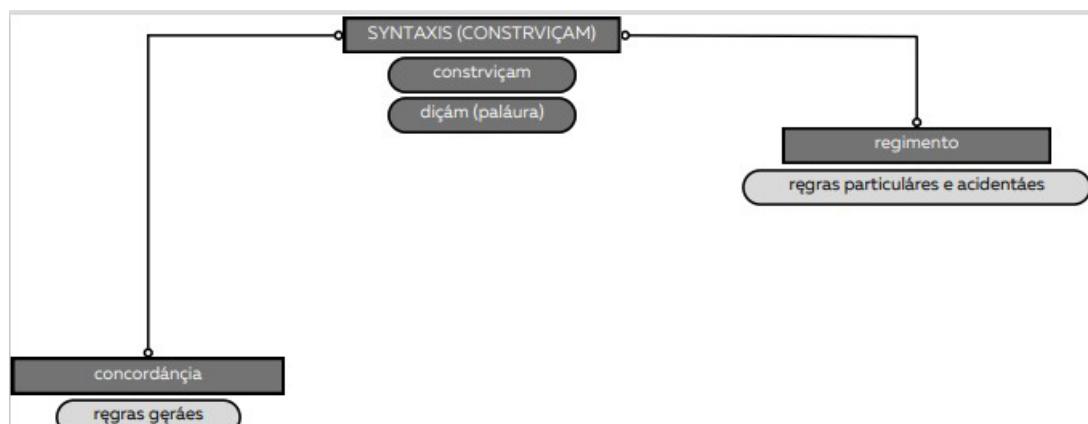
²² Nem todos os solecismos apresentados em Barros (1540) podem ser considerados, sob o olhar do linguista atual, como fenômenos ou desvios propriamente sintáticos. É o caso, por exemplo, da *ironia*, da *metáfora* e da *metonímia* (ver exemplificação e explicação desses metatermos na rede taxonômica e nos verbetes 37, 39 e 40 do glossário).

²³ Além desses 34 tipos de solecismo, no capítulo VII do Livro IV de Nebrija (1492, p. 50-54) há outros 16 tipos, ausentes em Barros (1540): *synthesis, acirologia, tapinosis, anadiplosis, anaphora, epanalepsis, paronomasia, omeoteleuton, omeoptoton, catachresis, periphrasis, hysterologia, anastropha, temesis, synchesis e calepos*.

3.2 Subdivisão da “syntaxis” e equivocidade do metatermo “construviçam”

A rede taxonômica da Figura 1 mostra que a bifurcação inicial do modelo sintático da *Grammatica* de Barros corresponde a “syntaxis” ramificada em “concordância” e “regimento”, conforme trecho da rede recortado a seguir:

Figura 2 – Divisão da sintaxe de Barros (1540) em concordância e regência



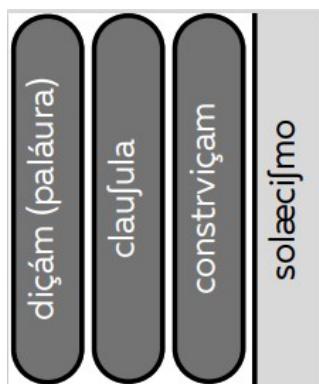
Fonte: Elaboração própria.

A Figura 2 evidencia não haver, no modelo sintático sob análise, a divisão primária entre “sintaxe simples/regular/natural” e “sintaxe figurada”, tão comum em gramáticas portuguesas subsequentes, especialmente no século 18, como nas obras de Argote (1725), Lobato (1770), Álvares (1786)²⁴, Figueiredo (1799) e Fonseca (1799). De modo estrito, portanto, a sintaxe em Barros (1540) equivale ao conjunto das regras de concordância e regência. Nesse modelo, as “figuras” – incluindo não apenas os “bárbarismos”, mas também os “solæcismos” – estão fora da sintaxe propriamente dita, tanto que são abordadas em capítulo próprio.

A Figura 2 ainda destaca que o metatermo “syntaxis” compartilha espaço com outro metatermo de mais alta frequência na obra: “construviçam”, oriundo da forma latina *construc-tio*, adaptada da forma grega *σύνταξις* [syntaxis]. Embora não seja de uso corrente na gramaticografia portuguesa subsequente (Schäfer-Prieß, 2019, p. 135), registra-se esse metatermo, também equivalendo à sintaxe, em Figueiredo (1799) e Fonseca (1799). Entretanto, em Barros (1540), há uma particularidade: “construviçam” é tanto sinônimo de “syntaxis” quanto objeto/unidade de análise da “syntaxis”, como revela a Figura 2, mas igualmente objeto/unidade de análise do “solæcismo”, como mostra este outro trecho recortado da rede:

²⁴ Álvares (1786, p. 35-44), embora adote essa mesma ideia, utiliza uma terminologia diferente para designar essa bifurcação inicial: subdivide a “sintaxe” em “das regras mais ordinárias” e “das figuras”.

Figura 3 – Objetos/unidades de análise do solecismo em Barros (1540)



Fonte: Elaboração própria.

Assim, a rede taxonômica representa essa peculiaridade terminológica do modelo sintático de Barros (1540) envolvendo o metatermo “construicam”. Por sua vez, os verbetes 20, 63 e 66 do glossário (ver Quadro 2) buscam organizar a multiplicidade de significados desse metatermos na obra, o qual é considerado ora como parte da gramática – “tratamos das primeiras tres pártes da grámatica. f. lettera, syllaba, diçã: fica agóra uermos a quàrta que é da cõstruicã [sic]” (Barros, 1540, p. 30) – ora como objeto/unidade de análise – “Ortografia, que tráta de letera, em Proſodia, que tráta de syllaba, em Ethimologia, que tráta da diçam e em Syntaxis, a que respõde a cõstruicã [sic]” (Barros, 1540, p. 2 verso).

Deve-se destacar, em contrapartida, que a equivocidade do metatermo “construicam” em Barros (1540) também é observada, de certo modo, no metatermo “sintaxe” na gramicografia ocidental dos séculos 17 a 19, utilizado para designar: a) as regras particulares a que cada língua recorre para combinar palavras em frases; b) a disciplina ou a parte da gramática responsável por descrever e explicar essas combinações; e c) fenômenos particulares no âmbito sintagmático (Auroux, 1979, p. 103; Fonseca, 2019, p. 163). Em última instância, pode-se afirmar que essas (e outras) interpretações decorrentes do uso de “sintaxe” e “construção” ainda vigoram no discurso contemporâneo da linguística quando se trata desse nível de análise gramatical.

3.3 Relações formais entre palavras e projeção da oração

Enquanto metatermo ou objeto/unidade de análise, a “oração” não desempenha papel relevante no modelo sintático de Barros (1540). Note-se a ausência desse metatermo na rede taxonômica e no glossário, embora ele esteja presente na *Grammatica* – mas sem qualquer definição – quando, por exemplo, se faz referência a “pártes da oraçam” (Barros, 1540, p. 12).

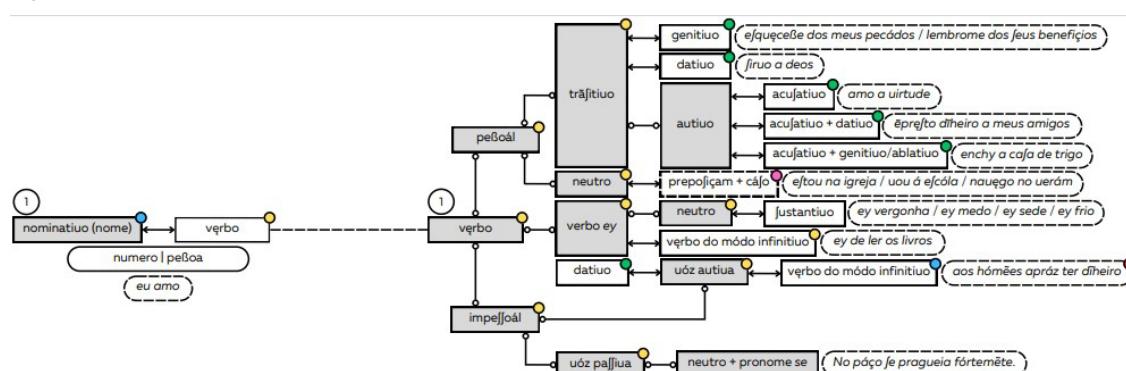
Isso se explica porque, dado que o modelo é estruturado a partir das regras de concordância e regência, a sintaxe na obra equivale às relações formais entre as *partes da oração*, ou seja, às relações formais entre as *palavras*. No domínio da “concordância”, essas relações são de identidade e envolvem os acidentes “numero”, “peſoa”, “gênero” e/ou “cáſo”; já no domínio do “regimento”, as relações são de determinação e envolvem, principalmente, o acidente

“cájo”. Na verdade, a unidade *palavra* é o ponto central de análise não apenas na sintaxe, mas em toda a *Grammatica* de Barros (Schäfer-Prieß, 2019, p. 316-317): a ortografia estuda suas letras; a prosódia, suas sílabas; a etimologia, suas formas.

Essa ideia de sintaxe circunscrita às relações formais entre palavras – a propósito, também presente em Nebrija (1492) – é introduzida na gramaticografia portuguesa, portanto, a partir de Barros (1540), percorrendo-a, com certa estabilidade, pelo menos até o final dos setecentos²⁵. Assim, durante esses três séculos, a estrutura da oração em português não foi descrita em sua completude lógico-semântica. Em vez disso, a ênfase analítica incidiu nas relações formais entre, de um lado, o nome no caso nominativo e o verbo (relação de concordância); e, de outro lado, o verbo e o nome nos demais casos (relação de regência).

A despeito dessa configuração, que pode ser visualizada na primeira parte da rede taxonômica (ver Figura 1), estruturas oracionais completas no português podem ser projetadas no modelo sintático de Barros (1540) a partir da consideração de relações lógicas não manifestas e metatermos latentes. Essas projeções são possíveis quando se nivelam a regra de “concordância” do tipo (1), centrada na relação formal de identidade entre o “nominatiuo/nome” e o “vêrbo”, à regra de “regimento” do tipo (1), ramificada nas relações formais de determinação entre diferentes tipos de “vêrbo” e seus respectivos “cásos”:

Figura 4 – Projeção da oração no modelo sintático de Barros (1540)



Fonte: Elaboração própria.

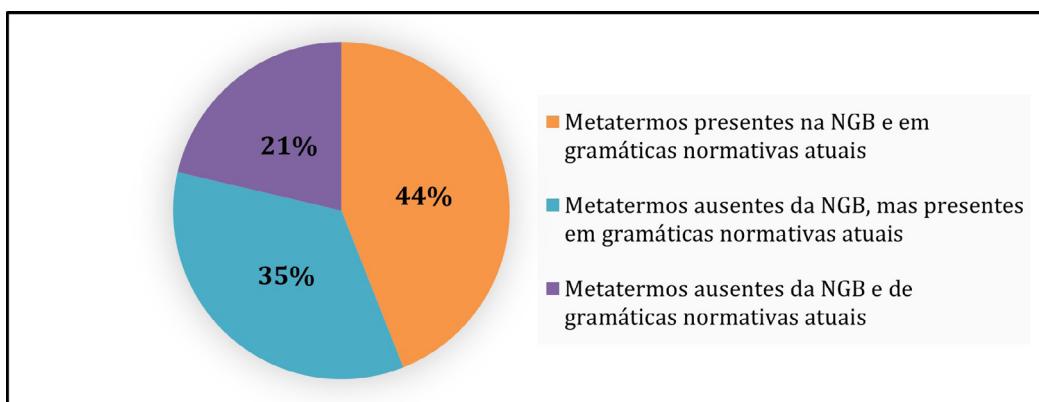
Dessa forma, na Figura 4, mais um trecho recortado da rede taxonômica apresentada na Figura 1, os círculos nas cores azul, amarelo, verde e rosa representam, respectivamente, as ideias de “sujeito”, “verbo”, “complemento verbal” e “adjunto adverbial” – embrionárias no modelo sintático dos quinhentos, mas fundamentais ao saber sintático contemporâneo no âmbito da gramaticografia (tradicional ou não), dos estudos linguísticos e das abordagens de ensino de gramática na educação básica.

²⁵ Entre os instrumentos linguísticos apresentados no Quadro 3 (1540-1799), a *Grammatica philosophica* de Mélo Bacellar (1783) é a única que se afasta, numa primeira análise, dessa macro-organização centrada nas relações de concordância e regência entre palavras, não apenas por apresentar efetivamente a “oração” (também chamada de “período” ou “proposição”) como objeto de análise sintática, mas também por realizar essa análise a partir da ideia de “partes essenciais da oração” (“agente”, “ação” e “acionado”), às quais se ligam “adjuntos” e “circunstâncias” (Mélo Bacellar, 1783, p. 13-14).

3.4 Metatermos quinhentistas e NGB

Embora metatermos estruturantes da oração na gramaticografia contemporânea – como “sujeito”, “complemento” e “adjunto” – não apareçam na *Grammatica* de Barros (1540), muito da terminologia das gramáticas elaboradas de acordo com a NGB (1959) já se apresenta no modelo sintático do autor:

Gráfico 1 – Presença da terminologia sintática de Barros (1540) na gramaticografia proveniente da NGB (1959)



Fonte: Elaboração própria.

O Gráfico 1 mostra que, dos 75 metatermos que compõem o modelo sintático quinhentista, 79% (44% + 35%) estão presentes na NGB e/ou em gramáticas normativas que a seguem, publicadas na segunda metade do século 20 e reeditadas, reimpressas e em circulação nos dias atuais, como Rocha Lima (2011), Bechara (2019), Mendes de Almeida (2019) ou Cunha e Cintra (2021). Além disso, o conteúdo (focal e contrastivo) desse grupo maior de metatermos na *Grammatica* de Barros é bem semelhante ao que se observa nessas obras contemporâneas. Destaque-se que os 21% restantes do total de metatermos se situam na segunda parte da rede taxonômica, dedicada aos solecismos. Isso significa que o conjunto de metatermos da “syntaxis” em Barros (1540) – a primeira parte da rede, ramificada em “concordância” e “regimento” – permanece integralmente na gramaticografia brasileira contemporânea de base tradicional, conforme detalha o Quadro 7:

Quadro 7 – Distribuição dos metatermos do modelo sintático de Barros (1540) nos grupos do Gráfico 1

75 metatermos		
33 presentes na NGB e em gramáticas normativas atuais	26 ausentes da NGB, mas presentes em gramáticas normativas atuais	16 ausentes da NGB e de gramáticas normativas atuais
aduérbio aietiuo appositiō cacophaton concordância eclypfis figura gēnero impeſſoál infinitiu interieçám módo numero paláura parenthēſis partícipio peſſoa peſſoál pleonaſmo prepoſiçam pronomē regimento relativo solæciſmo sylepſis syntaxis ſuſtantiuo trāſtituo uocatiuo uiçio verbo vóz autiuia vóz paſſiuia	ablatiuo acuſatiuo alegoria amphibologia anteçedēte antonomaſia appellatiuo autiuo cáſo clauſula construiçám datiuo epytheton genitiuo hyperbole metáfora metonomia neutro nome nominatiuo onomatopeia perioſſologia polyſyntheton prolepfis syncedoque zeuma	antiptoſis antyphraſis cacofyntheton charientiſmos dyaleton diçám enigma epizeuxis hirmos hypozenfis ironia macrologia paromeon polypteton scheſionomaton tantalogia

Fonte: Elaboração própria.

Expandindo a questão para além da sintaxe, Gómez Asencio e Gonçalves (2015) realizaram uma contagem aproximada dos metatermos presentes nas gramáticas de Nebrija e Barros que ainda estão vivos nas gramáticas atuais do espanhol e do português. Os pesquisadores também identificaram uma proporção consideravelmente maior de metatermos em uso em comparação aos metatermos obsoletos, o que ratifica o pioneirismo e a influência duradoura desses autores renascentistas. Vale lembrar, porém, que esse conjunto de metatermos, sobretudo os relacionados à sintaxe, não se desliga da cultura gramaticográfica greco-latina. Como Buescu (1984, p. 11) argumenta, as inovações terminológicas introduzidas por João de Barros no campo da sintaxe são de pouca monta quando comparadas aos feitos desse autor nos campos da morfologia e da fonética.

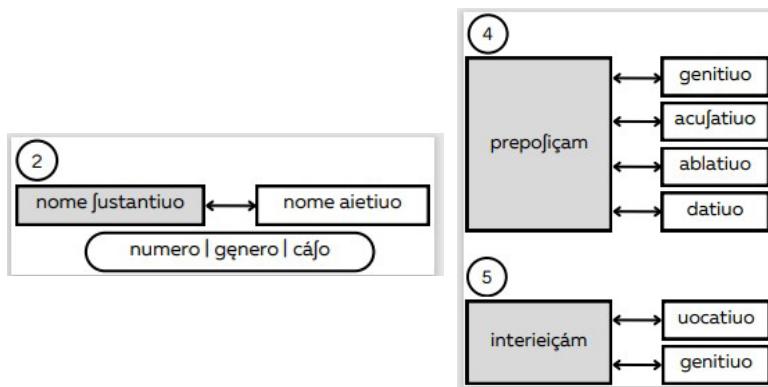
3.5 Sistema de casos e regência

Os esquemas gramaticais greco-latinos servem a Barros (1540) não apenas como fonte terminológica. O latim também aparece na obra como ponto de referência e modelo de codificação gramatical, sobretudo por razões pedagógicas. Na visão de Barros, o estudo da gramática portuguesa facilitaria o conhecimento do latim e de sua gramática. Assim, o autor enfatiza a aproximação dos fatos e sistemas gramaticais nas duas línguas (Buescu, 1984, p. 66-67; Borges Neto, 2022, p. 148-149). No campo da sintaxe, algumas das principais inovações da língua portuguesa em relação ao latim foram devidamente gramatizadas, como a existência do artigo e a formação perifrástica da voz passiva. Entretanto, o sistema de seis casos – “nominativo”, “genitivo”, “acusativo”, “dativo”, “ablativo” e “vocativo” – foi mantido. Nesse último aspecto, Barros (1540) foi seguido por quase todas as gramáticas de língua portuguesa vindouras até o aparecimento de Fonseca (1799), obra que se contrapõe explicitamente a características do que se entende por linhagem latinizada da gramaticografia portuguesa (cf. Faraco; Vieira, 2021, p. 486-489).

A rede taxonômica da Figura 1 revela que o metatermo “cáſo” integra o modelo sintático barroso de duas maneiras diferentes: explicitamente, como um dos acidentes da regra de “concordância” do tipo (2), na qual o “nome aietiuo” concorda em “cáſo” – além de

“gēnero” e “numero” – com o “nome Jstantiuo”; e implicitamente, como elemento determinado nas cinco regras de “regimento”, manifestando-se nos metatermos “genituo”, “acuſatiuo”, “datiuo”, “ablatiuo” ou “uocatiuo”. A Figura 5, ao recortar da rede taxonômica de Barros (1540) a regra (2) de “concordância” e as regras (4) e (5) de “regimento”, ilustra esses dois papéis do metatermo “cáſo” na obra:

Figura 5 – O metatermo “cáſo” como acidente da “concordância” e unidade do “regimento” em Barros (1540)



Fonte: Elaboração própria.

De toda sorte, ambos os arranjos decorrem da transposição de regras da gramática latina, e não de particularidades prescritivas ou características descritivas específicas do português. Esse desenho reforça a tese de que João de Barros “escreve uma gramática do português que reproduz, até onde é possível (e às vezes até onde não seria mais possível), a gramática latina” (Borges Neto, 2022, p. 149), ainda que enfatize a retórica “eles latinos vs. nós portugueses” face às diferenças incontornáveis entre as duas línguas.

A análise desses cinco tipos de “cáſo” nas regras de “regimento” exemplifica essa questão. Tais metatermos, centrais à gramaticografia latina, ganham em Barros (1540) significados mais amplos, muitas vezes imprecisos, associados a aspectos ontológicos e desvinculados das marcas morfológicas decorrentes das declinações latinas (ver verbetes 1, 2, 21, 30, 44 e 70 do glossário de metatermos), as quais não existem em português. Quando determinados pelo “regimento” do “vérbo”, alguns deles se aproximam, conforme já indicado na Figura 4, da ideia de “complemento verbal” da gramaticografia contemporânea²⁶. Em linhas gerais, diferentemente dos casos latinos, os casos portugueses apresentados em João de Barros não se distinguem pelas terminações, mas pela regência (Cardoso, 2004, p. 85; Borges Neto, 2022, p. 165), ou melhor, a partir da identificação da preposição e/ou do artigo que acompanha(m) o nome português.

3.6 O metatermo “regimento” e tipos de “vérbo peſoál”

Vale examinar mais detalhadamente a natureza do metatermo “regimento” [regência] e suas implicações na sintaxe de João de Barros, que assim explica esse tipo de relação: “Regimento

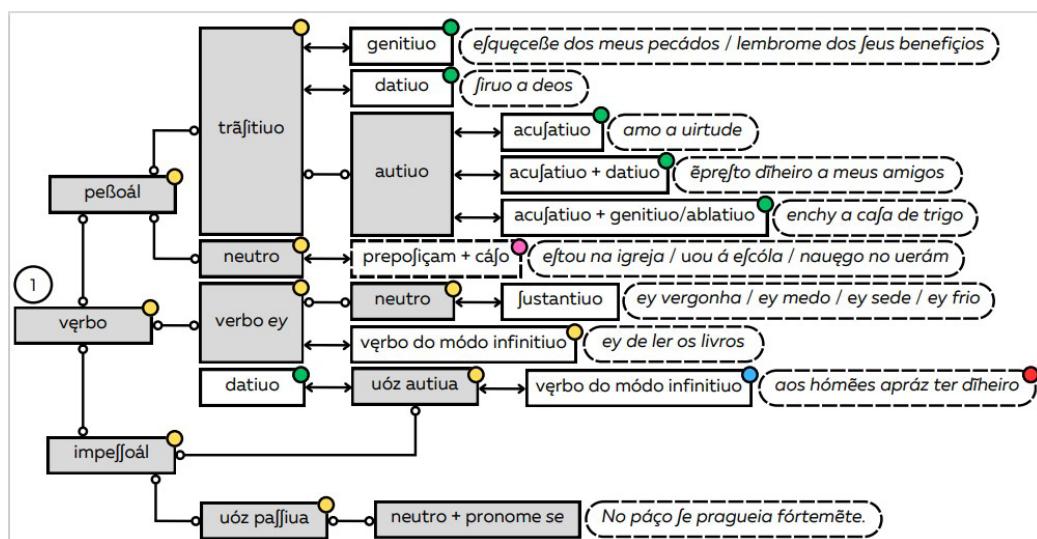
²⁶ A esse respeito, Moura Neves (2014, p. 38) afirma que “a explicitação da noção de regência vinculada determinantemente à noção de complementação é uma constante na tradição gramatical brasileira”.

é quando húa diçám se construe com outra diuerſa a ella, per gēnero ou per numero cáſo ou peſoa: sómente per húa espeſial natureza, cō que obriga e ſogeita a ſeguinte aſer pôſta em algum dos cáſos que temos" (Barros, 1540, p. 30 verso).

Nessa definição de "regimento", destacam-se duas características centrais à relação: a diferença morfológica e a dependência de caso entre as palavras. A rede taxonômica da Figura 1 mostra que a unidade que determina o caso (o termo regente) pode ser não apenas um "vérbo peſoál trāſituo", um "nome ſuſtantiuo/aietiuo" ou um "aduérbio" (como, em geral, des- crito atualmente), mas também uma "interieicám" ou "prepoſiçam". A inclusão dessa última unidade como elemento determinante da relação de "regimento", especificamente na regra do tipo (4), acarreta ao modelo sintático sobreposição descritiva, considerando-se as regras do tipo (1), (2) e (3), que apresentam unidades prepoſicionadas funcionando como elementos determinados nos casos "genitiuo", "datiuo" e "ablatiuo" (os termos regidos). Veja-se, por exemplo, diferentes configurações descritivas da construção "ſirvo a deos", presentes nas regras de "regimento" do tipo (1) ("vérbo peſoál trāſituo + datiuo") e do tipo (4) ("prepoſiçam + datiuo").

Outra sobreposição descritiva na relação de regência do modelo sintático de Barros (1540) envolve duas classificações distintas para o "vérbo peſoál". A primeira classificação é apresentada na terceira parte da gramática, que estuda a "diçam" (palavra), especificamente no capítulo "Do vérbo": trata-se da distinção entre "vérbo autiuo" e "vérbo neutro". A segunda classificação aparece na quarta parte da obra, dedicada à "syntaxis", exatamente no capítulo "Da constrviçam": refere-se ao par "vérbo trāſituo" e "vérbo neutro". É importante notar a recorrência do metatermo "neutro", ora em oposição ao metatermo "autiuo", ora em contraste com o metatermo "trāſituo". Essa situação singular pode ser observada na Figura 6, último recorte da rede taxonômica da sintaxe de Barros (1540):

Figura 6 – Tipos de "vérbo peſoál" no modelo sintático de Barros (1540)



Fonte: Elaboração própria.

A Figura 6 esclarece que, na sintaxe de Barros (1540), todo "vérbo autiuo" é um "vérbo trāſituo", mas nem todo "vérbo trāſituo" é um "vérbo autiuo", pois apenas este último rege exclusivamente unidades no "cáſo acuſatiuo". Por sua vez, o "vérbo neutro" se opõe ao "vérbo autiuo" porque "ſe nã pôde cõuerter ao módo paſſiuo" (BARROS, 1540, p. 18) e ao "vérbo

trāſitiuo” porque sua “auçā nā páſſa ē outra couſa” e “depois de ſy nam quērem cāſo ſenam mediāte prepoſiçā: como, Eſtou na igreia, uou á eſcólā” (p. 32). Essa primeira diferença enfatiza a dimensão estrutural da construção (ativa e passiva) como um todo, enquanto a segunda diferença considera, em certa medida, a dimensão lógico-relacional entre as palavras.

Também em certa medida, a oposição entre “vérbo trāſitiuo” e “vérbo neutro” aproxima, num horizonte de projeção, os conteúdos dos metatermos “neutro” e “intransitivo”, este último ausente do modelo sintático de Barros (1540). Na gramaticografia portuguesa, o metatermo “neutro” só começaria a ceder espaço para o metatermo “intransitivo” no final dos setecentos, em instrumentos como Fonseca (1799), cuja descrição gramatical tenta dar conta do “gênio da língua” – uma gramática de linhagem empirista, portanto (Faraco; Vieira, 2021, p. 486-489). Dado que, ao contrário do latim, não existem “verbos passivos” em português, não há razão para ser mantido o sistema verbal triádico latino “ativo”, “passivo” e “neutro”. Assim, o par “transitivo” e “intransitivo” se tornaria, numa sintaxe centrada em relações mais lógicas do que formais, uma alternativa mais apropriada.

Outra aproximação prospectiva que decorre do par “trāſitiuo” e “neutro” no modelo é a relação entre a unidade facultativa determinada pelo “vérbo neutro”, representada na Figura 6 como “prepoſiçām + cāſo”, e a ideia contemporânea de “adjunto adverbial”, na acepção de “termo acessório da oração”, como consta nas gramáticas que seguem a NGB. Observe-se que o elemento “prepoſiçām + cāſo” aparece sinalizado na rede por meio de um retângulo tracejado, que indica sua facultatividade, e de um círculo rosa, que representa essa ideia de “adjunto adverbial”.

Em suma, embora o “regimento” – ao lado da “concordância” – até hoje constitua território fértil para a inclusão de parâmetros avaliativos da construção sintática, inserindo-se privilegiadamente no chamado eixo da norma-padrão da gramática tradicional, na obra de Barros esse tema também se impõe não só pela possibilidade de obediência ao sistema de casos latinos (ao menos em terminologia), mas também pelo seu valor para o eixo da descrição metalinguística, prenunciando, ainda que muito timidamente, relações lógicas entre o verbo e outras partes da oração. Por meio do modelo sintático de Barros (1540), pode-se, assim, antever ideias linguísticas envolvendo predicação, transitividade, complementação, adjunção, entre outras a elas relacionadas e tão caras à descrição sintática contemporânea.

3.7 Construções complexas: coordenação e subordinação

Em Barros (1540), ou nas gramáticas portuguesas subsequentes, pelo menos até o século 18, não há análises sistematizadas de “construções complexas” (Moura Neves, 2016, p. 7-11), isto é, daquilo que a tradição gramaticográfica e pedagógica brasileira, à luz da NGB (1959), costuma chamar de “orações coordenadas e subordinadas” ou “período composto por coordenação e subordinação”. Isso explica a ausência desses processos sintáticos e de metatermos que os representem na rede taxonômica e no glossário aqui elaborados. Tal lacuna se deve, como esperado, à ênfase nas relações formais de concordância e regência, centradas nas palavras, em detrimento das relações lógicas entre as unidades constitutivas de construções (simples ou complexas) “dotadas de sentido completo”.

Por outro lado, na rede taxonômica, podem ser identificados exemplos que indicam a sensibilidade do gramático quinhentista em relação a construções complexas da língua,

embora lhe falte o aparato técnico e uma terminologia apropriada para sistematizar esses processos sintáticos. Assim, Barros (1540, p. 32) analisa a subordinação substantiva presente na construção “aos hómēes *apráz* ter dīheiro” (destacada na Figura 6 com um círculo vermelho) como decorrente da regência de verbos considerados impessoais, como “aprazer”, que “ante de *ſy* quérē dtō [dativo], e depois de *ſy* hū uérbo do módo infinito”. O autor apresenta exemplos análogos com outros verbos de classificação semelhante – “a my cōuē dár doutrina” [convir], “a ty *releua* aprēder ciença” [relevar] e “ás molheres cōpre onestidáde, e a todos obedecer aos preceitos da igreia” [cumprir] –, revelando e gramatizando a recorrência estrutural “datiuo + uérbo impeſſoál (uóz autiuia) + uérbo do modo infinitiuo” (ver Figura 6).

A maioria das construções complexas encontradas em Barros (1540), no entanto, são analisadas como espécies de “solæcīſmo”. Desse universo, na segunda parte da rede taxonômica (ver Figura 1), estão destacados com círculos coloridos exemplos de “figuras/uições” que se assemelham ou até mesmo equivalem a orações coordenadas – “prolepsīs”, “zeuma”, “hypozenſīs”, “sylepsīs” e “polyſyntheton” – ou subordinadas adjetivas – “antiptoſīs”, “macrologia”, “cacoſyntheton”, “hirmos” e “dyaleton”. Vale notar o uso do metatermo “clauſula” [cláusula] na explicação de alguns desses solecismos (ver verbetes 18, 23, 56 e 64 do glossário de metatermos), o que sugere que essas construções envolvem mais do que relações entre palavras ou partes da oração.

Por fim, é relevante mencionar que Barros estabelece duas espécies de “coniunçám” [conjunção]: a “copulatiua” [copulativa] e a “diſiuntiua” [disjuntiva], exemplificadas respectivamente pelos elementos “e” e “ou”. Entretanto, a abordagem dessas unidades não mobiliza estruturas oracionais coordenadas, limitando-se à coordenação de palavras, como, por exemplo, em “Alexādre e Cēſar e Hanibál e Pōpēo e Pirro, forā grandes capitāes” (Barros, 1540, p. 33 verso). Desse modo, por não fazerem parte da macro-organização ou das relações internas do modelo sintático de Barros (1540), esses três metatermos não foram apresentados na rede taxonômica nem no glossário.

4 Últimas palavras

Este artigo apresentou duas ferramentas autorais bastante produtivas para análise, reconstrução, sistematização e comparação de modelos sintáticos explícitos ou subjacentes na história da gramaticografia ocidental: a *rede taxonômica* e o *glossário de metatermos*. Isso foi feito a partir do escrutínio das ideias sobre sintaxe contidas na *Grammatica da língua Portuguesa* (1540), de João de Barros.

A análise terminográfica desenvolvida foi do tipo “monográfica/isotópica” (Swiggers, 2009b, p. 24), pois se concentrou na terminologia sintática de uma obra específica de um autor específico. Em trabalhos futuros, de posse desses resultados, pretendo desenvolver terminografias do tipo “comparativa/contrastiva”. Isso envolverá o cotejo e a contraposição entre o modelo sintático de Barros (1540) e os modelos sintáticos dos instrumentos linguísticos listados no Quadro 3: Roboredo (1619), Gama/Argote (1721/1725), Cunha (1788 [1769]), Lobato (1770, 1797), Bacellar (1783), Álvares (1786), Casimiro (1792), Figueiredo (1799) e Fonseca (1799).

O objetivo, a médio prazo, é elaborar uma *síntese de alcance hermenêutico*²⁷ da sintaxe desenvolvida nos três primeiros séculos de gramaticografia portuguesa, prenúncio da gramaticografia brasileira. Trata-se de uma empreitada historiográfica necessária e, até o momento, inédita, que abrange a criação de uma história longitudinal dos modelos e da terminologia sintática luso-brasileira, bem como a produção de um dicionário da evolução dos metatermos que os constituem.

Referências

- ALTMAN, C. História, estórias e historiografia da linguística brasileira. *Todas as Letras*, São Paulo, v. 14, n. 1, 2012. p. 14-37. Disponível em: <<http://editorarevistas.mackenzie.br/index.php/tl/article/view/4526>>. Acesso em: 02 nov. 2023.
- ÁLVARES, F. de C. *Breve compendio da gramatica portugueza para o uso Das Meninas que se educaõ no Mosteiro da Vizitaçāo de Lisboa*. Lisboa: Officina de Antonio Rodrigues Galhardo, 1786.
- ARCOTE, J. C. de. *Regras da lingua Portugueza, espelho da lingua Latina, Ou dispoñiçaõ para facilitar o ensino de lingua Latina pelas regras da Portugueza*. 2^a. impressão. Lisboa: Officina da Musica, 1725 [1721].
- ARISTÓTELES. *Categorias*. Trad. Ricardo Santos. Porto: Porto Editora, 1995.
- ARNAULD; LANCELOT. *Gramática de Port-Royal ou Gramática geral e razoada*. São Paulo: Martins Fontes, 1992 [1660].
- AUROUX, S. *La sémiotique des encyclopédistes*. Paris: Payot, 1979.
- AUROUX, S. *A revolução tecnológica da gramatização*. 3^a ed. Trad. E. P. Orlandi. Campinas, SP: Editora da Unicamp, 2014 [1992].
- BACELLAR, B. de L. e M. *Grammatica philosophica, e orthographia racional da Lingua Portugueza; Para se pronunciarem, e escreverem com acerto os vocabulos desse idiôma*. Lisboa: Offic. de Simão Thaddeo Ferreira, 1783.
- BARROS, D. L. P. de. O discurso da gramática do português. *Revista da ABRALIN*, v. Eletrônico, n. Especial, p. 291-332, 1^a parte 2011. Disponível em: <<https://revista.abralin.org/index.php/abralin/article/view/1094>>. Acesso em: 02 nov. 2023.
- BARROS, J. D. A história Serial e história Quantitativa no movimento dos Annales. *História Revista*, Goiânia, v. 17, n. 1, p. 203-222, jan.-jun. 2012. DOI: 10.5216/hr.v17i1.21693.
- BARROS, J. de. *Grammatica da lingua portuguesa com os mandamentos da Santa madre igreja*. Lisboa: Casa de Luís Rodriguez, 1539.
- BARROS, J. de. *Grammatica da lingua Portuguesa Dialogo em lovvor da nossa lingvagem*. Lisboa: Olyssippone. Apud Lodouicum Rotorigū, 1540.

²⁷ Koerner (1996a, p. 47) afirma, nessa direção, que a prática historiográfica requer capacidade de síntese, isto é, a faculdade de destilar o essencial da massa dos fatos empíricos coligidos a partir das fontes primárias, a fim de interpretar as descobertas e lhes fornecer uma explicação adequada.

- BATISTA, R. de O. *Fundamentos da pesquisa em Historiografia da Linguística*. São Paulo: Editora Mackenzie, 2020.
- BECCARI, A. J. *Tratado sobre os modos de significar da gramática especulativa, de Tomás de Erfurt*. Curitiba: Editora da UFPR, 2019.
- BECHARA, E. *Moderna gramática portuguesa*. 39^a ed. Nova ed. rev. e ampl. pelo autor. Rio de Janeiro: Nova Fronteira, 2019 [1961].
- BECKER, C. L. *The Heavenly City of the Eighteenth-Century Philosophers*. New Haven; London: Yale Univ. Press, 1932.
- BOLIQUEIME, E. M. da C. *As figuras de construção na gramaticografia portuguesa do século XVIII*. 2021. 136 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade do Porto, Porto, 2021.
- BORGES NETO, J. *História da gramática*. Curitiba: Editora da UFPR, 2022.
- BOTAS, V. B. Introducción. In: DÍSCOLO, A. *Sintaxis*. Trad. Vicente Bécares Botas. Madrid: Editorial Gredos, 1987. p. 9-70.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO E CULTURA. *Nomenclatura Gramatical Brasileira*. Diário Oficial de 11 mai. 1959.
- BUESCU, M. L. C. *Historiografia da língua portuguesa: século XVI*. Lisboa: Livraria Sá da Costa Editora, 1984.
- CARDOSO, S. O estudo da palavra na gramática portuguesa no séc. XVI. In: BRITO, A. M. (org.). *Linguística histórica e história da língua portuguesa: actas do Encontro de Homenagem a Maria Helena Paiva*. Porto: Universidade do Porto; Faculdade de Letras, 2004, p. 73-87.
- CASAGRANDE, N. do S. *A implantação da língua portuguesa no Brasil do Século XVI: um percurso histórico-gramatical*. São Paulo: EDUC, 2005.
- CASIMIRO, J. J. *Methodo grammatical resumido da lingua portugueza*. Porto: Offic. de Antonio Alvarez Ribeiro, 1792.
- CAVALIERE, R. *A gramática no Brasil: ideias, percursos e parâmetros*. Rio de Janeiro: Lexikon, 2014.
- CAVALIERE, R. *História da gramática no Brasil: séculos XVI a XIX*. Petrópolis: Vozes, 2022.
- COLOMBAT, B.; FOURNIER, J-M.; PUECH, C. *Uma história das ideias linguísticas*. Trad. Jacqueline Léon e Marli Quadros Leite. São Paulo: Contexto, 2017.
- CUNHA, C.; CINTRA, L. F. L. *Nova gramática do português contemporâneo*. 7^a ed. Rio de Janeiro: Lexikon, 2021 [1985].
- CUNHA, J. P. F. da. *Breve tratado da orthografia para os que não frequentaraõ os estudos [...]*. 6^a impressão. Officina de Antonio Gomes, 1788 [1769].
- D'OLIVEIRA, F. *Grammatica da lingoagem portuguesa*. Lisboa: Casa d'Germão Galharde, 1536.
- D'OLIVEIRA, F. *Grammatica da linguagem portugueza*. 2^a ed., conforme a de 1536. Imprensa Portuguesa, 1871 [1536].
- DIAS, L. F. Articulação sintática em gramáticas do século XIX. *Letras*, Santa Maria, RS, vol. 18, n. 2, p. 125-134, 2008. DOI: <https://doi.org/10.5902/2176148511983>.

FARACO, C. A.; VIEIRA, F. E. A linhagem empirista na gramaticografia do século 18. *Revista da Abralin*, v. 20, p. 464-492, 2021. DOI: 10.25189/rabralin.v20i3.1963.

FERNANDES, G. Os *Grammatices Rudimenta* (>1540) de João de Barros. *Boletim de Estudos Clássicos*, Associação Portuguesa de Estudos Clássicos / Instituto de Estudos Clássicos da Universidade de Coimbra, n. 43, p. 131-136, jun. 2005. Disponível em: <<https://repositorio.utad.pt/handle/10348/6115>>. Acesso em: 25 ago. 2023.

FERREIRA, E. G. de M. *Uma historiografia do processo brasileiro de gramatização da colocação pronominal em gramáticas oitocentistas*. 2021. 232 f. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal da Paraíba, João Pessoa, 2021.

FIGUEIREDO, P. J. de. *Arte da grammatica portugueza ordenada em methodo breve, facil, e claro*. Lisboa: Regia Officina Typografica, 1799.

FONSECA, M. do C. Conceitos-chave do discurso historiográfico português sobre a sintaxe. In: TEIXEIRA, M. (org.). *Estudos da língua portuguesa: a união na diversidade*. Santarém: Instituto Politécnico de Santarém; Escola Superior de Educação, 2019. p. 99-129.

FONSECA, P. J. da. *Rudimentos da grammatica portugueza, Cómmodos á instrucçāo da Mocidade, em confirmados com selectos exemplos de bons Autores*. Lisboa: Off. de Simão Thaddeo Ferreira, 1799.

GAMA, C. M. da. *Regras da lingua Portugueza, espelho da lingua Latina, Ou disposiçām Para facilitar o ensino de lingua Latina pelas regras da Portugueza*. Lisboa: Officina de Mathias Pereyra da Sylva & Joaõ Antunes Pedrozo, 1721.

GÂNDAVO, P. de M. de. *Regras que ensinam a maneira de escrever e a ortografia da língua portuguesa*. Edição fac-similada da 1ª ed. Introdução de Maria Leonor Carvalhão Buescu. Lisboa: Biblioteca Nacional, 1981 [1574].

GÓMEZ ASENCIO, J. J.; MONTORO DEL ARCO, E. T. M.; SWIGGERS, P. *Principios, tareas, métodos e instrumentos en historiografía lingüística*. In: VAQUERA, M. L. C. et al. (Eds.). *Métodos y resultados actuales en historiografía de la lingüística*. Nodus Publikationen, 2014. p. 266-301.

GÓMEZ ASENCIO, J.; GONÇALVES, M. F. Terminologia gramatical luso-castelhana dos inícios: de Antonio de Nebrija e João de Barros. *Confluência*, Rio de Janeiro, n. 49, v. 2, p. 68-118, jul-dez. 2015. DOI: <https://doi.org/10.18364/rc.v1i49.80>.

KEMMLER, R. Caetano Maldonado da Gama, D. Jerónimo Contador de Argote e as duas edições das *Regras da lingua portugueza, espelho da lingua latina* (1721, 1725). *Limite*, n. 6, p. 75-101, 2012. Disponível em: <<https://www.revistalimite.es/volumen%206/05kemm.pdf>>. Acesso em: 23 out. 2023.

KEMMLER, R.; ASSUNÇÃO, C.; FERNANDES, G. A primeira gramática portuguesa para o ensino feminino em Portugal (Lisboa, 1786). *Diacrítica, Revista do Centro de Estudos Humanísticos*, Universidade do Minho, v. 24, n. 1, p. 373-393, 2010. Disponível em: <https://www.researchgate.net/publication/269037968_A_primeira_gramatica_portuguesa_para_o_ensino_feminino_em_Portugal_Lisboa_1786>. Acesso em: 01 nov. 2023.

KOERNER, K. Questões que persistem em historiografia linguística. *Revista da ANPOLL*, n. 2, p. 45-70, 1996a. DOI: <https://doi.org/10.18309/anp.v1i2.240>.

KOERNER, E. F. K. O problema da metalínguagem em historiografia da linguística. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 12, n. 1, p. 95-124, 1996b. Disponível em: <<https://revistas.pucsp.br/index.php/delta/article/view/43754>>. Acesso em: 03 nov. 2023.

- KOERNER, E. F. K. A importância da historiografia linguística e o lugar da história nas ciências da linguagem. In: KOERNER, E. F. K. *Quatro décadas de historiografia linguística: estudos selecionados*. Trás-os-Montes e Alto Douro: Centro de Estudos em Letras, Universidade de Trás-os-Montes e Alto Douro, 2014a. p. 9-15.
- KOERNER, E. F. K. O problema da metalinguagem em historiografia da linguística. In: KOERNER, E. F. K. *Quatro décadas de historiografia linguística: estudos selecionados*. Trás-os-Montes e Alto Douro: Centro de Estudos em Letras, Universidade de Trás-os-Montes e Alto Douro, 2014b. p. 75-90.
- KOERNER, E. F. K. The historiography of linguistics past, present, future. In: KOERNER, E. F. K. *Last Papers in Linguistic Historiography*. John Benjamins Publishing Company, 2020. p. 3-35.
- LAUDAN, L. *O progresso e seus problemas: rumo a uma teoria do crescimento científico*. Trad. R. L. Ferreira. São Paulo: Editora Unesp, 2011 [1978].
- LEITE, M. Q. *O nascimento da gramática portuguesa: uso & norma*. São Paulo: Humanitas; Paulistana, 2007.
- LEITE, M. Q. A gramatização da colocação dos pronomes átonos em gramáticas portuguesas e brasileiras. *Todas as Letras*, São Paulo, v. 15, n. 2, p. 127-140, 2013. Disponível em: <<http://editorarevistas.mackenzie.br/index.php/tl/article/view/5617>>. Acesso em: 05 abr. 2022.
- LIÃO, D. N. de. *Orthographia da lingoa Portuguesa*. Lisboa: Antonio Gonsalves, 1576.
- LIMA, L. C. de. *Orthographia da Lingua Portugueza*. Lisboa: Officina de Antonio Isidoro da Fonseca, 1736.
- LOBATO, A. J. dos R. *Arte da grammatica da lingua portugueza [...]*. Lisboa: Regia Officina Typografica, 1770.
- LOBATO, A. J. dos R. *Arte da grammatica da lingua portugueza [...]*. 4^a impressão. Lisboa: Regia Officina Typografica, 1797 [1770].
- MENDES DE ALMEIDA, N. *Gramática metódica da língua portuguesa*. 46^a ed. São Paulo: Saraiva, 2019 [1944].
- MORAES SILVA, A. de. *Epitome da grammatica da lingua portugueza*. Lisboa: Off. de Simão Thaddeo Ferreira, 1806.
- MOURA NEVES, M. H. de. Uma avaliação do conceito de regência e transitividade na tradição gramatical do português. *Todas as Letras*, São Paulo, v. 16, n. 1, p. 34-47, mai. 2014. Disponível em: <<https://editorarevistas.mackenzie.br/index.php/tl/article/view/6762>>. Acesso em: 30 out. 2023.
- MOURA NEVES, M. H. de. (org.). *A construção das orações complexas*. Ataliba T. de Castilho (Coord.). Gramática do português culto falado no Brasil, v. 5. São Paulo: Contexto, 2016.
- NEBRIJA, E. A. de. *Gramática sobre la lengua castellana*. Edición de Miguel Ángel Esparza y Ramón Sarmiento. Madrid: Fundación Antonio de Nebrija-SGEL, 1992 [1492].
- POLACHINI, B. S. Considerações sobre o impacto da *Grammaire Générale et Raisonnée de Port-Royal* (1660) no tratamento da sintaxe de gramáticas brasileiras do português do século XIX. *Confluência*, Rio de Janeiro, n. 39-40, p. 296-314, jul.-dez. 2010, jan.-jun. 2011. Disponível em: <<https://revistaconfluencia.org.br/rc/article/view/673>>. Acesso em: 02 nov. 2023.
- POLACHINI, B. S. *Uma história serial e conceitual da gramática brasileira oitocentista de língua portuguesa*. 2018, 458 f. Tese (Doutorado em Letras), Programa de Pós-Graduação em Semiótica e Linguística Geral, Universidade de São Paulo, São Paulo, 2018.
- POLACHINI, B. Análise de textos e metatextos gramaticais: Costa Duarte (1829, 1853, 1859, 1877) e Bithencourt (1862) sobre o 'verbo substantivo'. *Cadernos de Pós-Graduação em Letras (CPGL)*, São Paulo,

v. 16, p. 17-30, 2016. Disponível em: <<https://editorarevistas.mackenzie.br/index.php/cpgl/article/view/9527>>. Acesso em 23 out. 2023.

POLACHINI, B. S.; DANNA, S. M. D. G. Dados sintáticos do português brasileiro em gramáticas brasileiras oitocentistas. *Estudos Linguísticos*, São Paulo, vol. 45, n. 1, p. 192-202, 2016. DOI: <https://doi.org/10.21165/el.v45i1.800>.

PRISCIEN. *Grammaire, Livre XVII – Syntaxe*, 1. Trad. Groupe Ars Grammatica. Paris: Librairie Philosophique J. Vrin, 2010.

ROBINS, R. H. *Pequena história da linguística*. Trad. L. M. M. de Barros. Rio de Janeiro: Ao Livro Técnico; Brasília: INL, 1979.

ROBOREDO, A. de. *Methodo grammatical para todas as lingvas*. Lisboa: Pedro Craesbeeck / The University of Chicago Library, 1619

ROCHA LIMA, C. H. da. *Gramática normativa da língua portuguesa*. 49ª ed. Rio de Janeiro: José Olympio, 2011 [1957].

SCHÄFER-PRIEß, B. A *Gramaticografia Portuguesa até 1822*: condições da sua gênese e critérios de categorização, no âmbito da tradição latina, espanhola e francesa. Trad. J. F. da Silva. Vila Real, Portugal: Centro de Estudos em Letras / Universidade de Trás-os-Montes e Alto Douro, 2019. Disponível em: <https://www.utad.pt/cel/wp-content/uploads/sites/7/2020/10/Lingui%CC%81stica_14.pdf>. Acesso em: 01. mar. 2023.

SILVA, M. Princípios metodológicos e fundamentação teórica da gramaticografia - por uma história cultural da gramática portuguesa. *Revista da ABRALIN*, v. 5, n. 1 e 2, p. 61-81, dez. 2006. Disponível em: <<https://revista.abralin.org/index.php/abralin/article/view/940/867>>. Acesso em: 02 nov. 2023.

SWIGGERS, P. Modelos, métodos y problemas en la historiografía de la Lingüística. In: ZUMBADO, C. et al. (Eds.). *Nuevas aportaciones a la Historiografía Lingüística*. In: Congresso Internacional de Lasehl, v. 4, Actas... Madrid: Arco Libros, 2004. p. 113-146.

SWIGGERS, P. Terminologie et terminographie linguistiques: problèmes de définition et de calibrage. *Syntaxe & Sémantique – La terminologie linguistique*, n. 7. Caen: Presses universitaires de Caen, p. 13-28, 2006. Disponível em: <<https://www.cairn.info/revue-syntaxe-et-semantique-2006-1-page-13.htm>>. Acesso em 19 out. 2023.

SWIGGERS, P. La historiografía de la lingüística: apuntes y reflexiones. *Revista Argentina de Historiografía Lingüística*. Buenos Aires, v. 1, n. 1, p. 67-76, 2009a. Disponível em: <<https://www.rahl.ar/index.php/rahl/article/view/6>>. Acesso em: 02 nov. 2023.

SWIGGERS, P. Terminología gramatical y lingüística: elementos de análisis historiográfico y metodológico. *Res Diachroniae*, v. 7, 2009b, p. 11-35. Disponível em: <https://resdiachroniae.files.wordpress.com/2013/12/volumen-7-04_swiggers_pierre.pdf>. Acesso em: 29 out. 2023.

SWIGGERS, P. História e Historiografia da Linguística: *status*, modelos e classificações. *Eutomia: Revista de Literatura e Linguística*, Recife, ano III, v. 2, dez. 2010a. p. 1-17. Disponível em: <<https://periodicos.ufpe.br/revistas/index.php/EUTOMIA/article/view/1702>>. Acesso em: 02 nov. 2023.

SWIGGERS, P. Le métalangage de la linguistique: réflexions à propos de la terminologie et de la terminographie linguistiques. *Revista do CEL*, São Paulo, v. 7, n. 2, p. 9-29, 2010b. Disponível em: <<https://revistas.gel.org.br/rg/article/view/61>>. Acesso em: 14 out. 2023.

- SWIGGERS, P. Linguistic historiography: object, methodology, modelization. *Todas as Letras*, São Paulo, v. 14, n. 1, p. 38-53, 2012. Disponível em: <<https://editorarevistas.mackenzie.br/index.php/tl/article/view/4527>>. Acesso em: 24 out. 2023.
- SWIGGERS, P. A historiografia da linguística: objeto, objetivos, organização. *Confluência*, n. 44-45, 2013. p. 39-59. Disponível em: <<https://revistaconfluencia.org.br/rc/article/view/602>>. Acesso em: 02 nov. 2023.
- SWIGGERS, P. Directions for linguistic historiography. *Cadernos de Historiografia Linguística do CEDOCH*: VII MiniEnapol de Historiografia Linguística (2013). São Paulo, v. 1, p. 8-17, 2015. Disponível em: <https://cedoch.fflch.usp.br/sites/cedoch.fflch.usp.br/files/u65/CHLC1_o.pdf>. Acesso em: 02 nov. 2023.
- SWIGGERS, P. Gramaticografía e historiografía: una visión retro- y prospectiva. *Anales de Lingüística – Segunda época*, Mendoza, Argentina, n. 4, p. 139-154, abr.-set. 2020. Disponível em: <<https://revistas.uncu.edu.ar/ojs3/index.php/analeslinguistica/article/view/4393>>. Acesso em: 02 nov. 2023.
- VIEIRA, F. E. *A gramática tradicional*: história crítica. São Paulo: Parábola Editorial, 2018.
- VIEIRA, F. E. A sintaxe no Brasil: notas historiográficas e eixos temáticos de investigação. *Alfa: Revista de Linguística*, São José do Rio Preto, SP, v. 64, p. 1-29, 2020a. DOI: <https://doi.org/10.1590/1981-5794-e12288>.
- VIEIRA, F. E. Tradição gramatical: história, epistemologia e ensino. In: VIEIRA, F. E.; BAGNO, M. (orgs.). *História das línguas, histórias da Linguística*: homenagem a Carlos Alberto Faraco. São Paulo: Parábola Editorial, 2020b. p. 85-124.
- VIEIRA, F. E.; FARACO, C. A. *Gramática do português brasileiro escrito*. São Paulo: Parábola Editorial, 2023.

Construção de *corpus* de 80 verbos de causalidade implícita no Português Brasileiro

Construction of a Corpus of 80 Implicit Causality Verbs in Brazilian Portuguese

Rute da Silva Barbalho
Universidade Federal do Rio Grande do Norte (UFRN) | Natal | RN | BR
rute.barbalho.014@ufrn.edu.br
<https://orcid.org/0000-0002-5912-5321>

Renata Sabrinne Souza de Carvalho
Universidade Federal do Rio Grande do Norte (UFRN) | Natal | RN | BR
professorarenatasabrinne@gmail.com
<https://orcid.org/0000-0001-5254-6476>

Mahayana Cristina Godoy
Universidade Federal do Rio Grande do Norte (UFRN) | Natal | RN | BR
mahayana.godoy@ufrn.br
<https://orcid.org/0000-0002-7499-3290>

Resumo: O presente estudo tem por objetivo construir um *corpus* com verbos de viés de causalidade implícita, para ser usado posteriormente nos mais diversos tipos de trabalho. A Causalidade implícita é uma propriedade semântica que está presente em alguns verbos, os quais têm tendência a retomar a causa de determinado evento (Goikoetxea, Acha e Pascual, 2008). Na oração “Maria admirava João”, por exemplo, “admirava” é um verbo de causalidade implícita pois aponta para a causa da admiração, “João”, em detrimento de “Maria”. Assim, a partir de um experimento *offline* de complementação de sentenças, testamos qual o viés de CI de 80 verbos no PB (40 de viés SN1 e 40 de viés SN2), os quais são divididos em quatro categorias de acordo com a sua taxonomia: Agente-Paciente (AP), Agente-Evocador (AE), Estímulo-Experienciador (SE) e Experienciador-Estímulo (ES) (Rudolph, 1997). Além disso, verificamos se o gênero do participante da pesquisa influenciava o estabelecimento da correferência (Ferstl, Garnham e Manouilidou, 2011), como também o gênero gramatical dos participantes das sentenças, tendo em vista que homens são percebidos como mais causais que as mulheres (LaFrance, Brownell e Hahn, 1997). Os nossos resultados confirmam o viés de causalidade dos verbos: para AP e SE houve uma preferência de direcionar a causa para o SN1, enquanto em AE e ES há uma tendência de direcionamento da causa para o SN2, porém, não encontramos evidências da relação entre o gênero do participante e a CI.

Palavras-chave: causalidade implícita; *corpus*; verbos; português brasileiro.



Abstract: The present study aims to build a corpus with implicit causality bias verbs, to be used later in the most diverse types of work, based on constraint-based models. Implicit Causality is a semantic property that is present in some verbs that tend to refer to the cause of a certain event (Goikoetxea, Acha and Pascual, 2008). In the sentence “Maria admired João”, for example, “admired” is an implicit causality bias verb, because it points to the cause of admiration, “João”, to the detriment of “Maria”. Thus, from an offline sentence completion experiment, we tested the IC bias of 80 verbs in brasilian portuguese (40 with SN1 bias and 40 with SN2 bias), which are divided into four categories according to their taxonomy: Agent-Patient (AP), Agent-Evocator (AE), Stimulus-Experiencer (SE) and Experiencer-Stimulus (ES). Furthermore, we verified whether the gender of the research participant influenced the establishment of coreference (Ferstl, Garnham and Manouilidou, 2011), as well as the grammatical gender of the sentence participants, considering that men are perceived as more casual than women (LaFrance, Brownell and Hahn (1997). Our results confirm the causality bias of verbs: for AP and SE, there was a preference for assigning the cause to SN1, while in AE and ES, there is a tendency to assign the cause to SN2. However, we didn't find evidence of a relationship between the participant's gender and IC bias.

Keywords: implicit causality; corpus; verbs; Brazilian Portuguese.

1 Introdução

Nas nossas interlocuções, algumas vezes não apenas comunicamos informações explícitas, mas também podemos indicar relações de causa de modo implícito. Essas indicações de causa subjacentes entre eventos, ideias ou ações podem ser percebidas por meio do contexto, das escolhas vocabulares e/ou da estrutura das sentenças utilizadas pelos falantes.

Uma dessas indicações de causa pode ocorrer em contextos em que o fenômeno da Causalidade Implícita se apresenta. A Causalidade Implícita (doravante CI) é um fenômeno linguístico manifestado no conteúdo semântico de determinadas formas verbais que direciona a atribuição da causa de um determinado evento descrito por um verbo interpessoal (evento/ação que envolve dois participantes) (Goikoetxea;Pascual; Acha, 2008, p.760-761). Vejamos os exemplos a seguir:

- (1) João entediava Maria.
- (2) Maria admirava João.

Nas sentenças apresentadas acima as formas verbais utilizadas podem ser classificadas como verbos de causalidade implícita. Logo, esses verbos possuem um viés causal que é interpretado de maneira inconsciente. Esse viés causal pode ora atribuir a causa para o evento descrito pelo verbo para o sujeito da sentença (sintagma nominal 1 – SN1) ora para o objeto da sentença (sintagma nominal 2 – SN2). No caso da sentença (1), “entediava” é uma condição em que a causa comumente evoca o sujeito da oração (João), assim, de forma implícita, a ação de entediar apresenta uma causa que só poderia ser atribuída por aquele que provoca o tédio. Já na sentença (2) podemos observar um viés contrário, já que “admirava” geralmente alude à causa da admiração que está compreendida no objeto da oração (João).

Desse modo, os verbos de Causalidade Implícita podem ser divididos a partir das suas duas tendências: determinado verbo que apresenta uma inferência de atribuição causal ao sujeito é comumente denominado de verbo com tendência para o SN1; já os verbos cuja inferência causal é direcionada ao objeto, geralmente, são referidos como verbos com tendência para o SN2 (Goikoetxea; Pascual; Acha, 2008). A partir disso, poderíamos classificar os verbos do exemplo anterior da seguinte forma: “entediar” – verbo com tendência para o SN1; “admirar” – verbo com tendência para o SN2. Essas tendências são passíveis de verificação, isto é, podemos investigar se o fenômeno da CI ocorre efetivamente nos verbos interpessoais de diversas línguas do mundo.

Nesse sentido, uma vez que compreendemos do que se trata a causalidade implícita, temos por objetivo construir um *corpus* desses verbos no PB, que poderá ser usado em pesquisas futuras que investiguem como falantes dessa língua processam informação implícita. Para isso, começamos este trabalho com uma seção introdutória dividida em três subseções de referencial teórico: uma revisão sobre estudos de causalidade implícita em diversas línguas, a apresentação de uma taxonomia dos verbos de causalidade implícita segundo Rudolph (1997) e, por fim, uma revisão sobre a interação entre causalidade implícita e o gênero dos participantes do evento e dos sujeitos que participam em pesquisas psicolinguísticas. A partir dessas discussões, justificamos o experimento de continuação de sentenças que apresentamos nas seções de metodologia para construir o *corpus* de CI em PB.

1.1 Estudos sobre Causalidade Implícita em diferentes idiomas

As tendências que descrevemos com os exemplos (1) e (2) podem ser verificadas, especialmente, por meio de pesquisas empíricas. A partir do levantamento histórico desenvolvido por Goikoetxea, Acha e Pascual (2008) podemos notar a abrangência das pesquisas sobre a Causalidade Implícita. Segundo os autores, a partir da primeira pesquisa de caráter empírico que corresponde ao trabalho de Garvey e Caramazza (1974), a pesquisa sobre a CI passou a ser contemplada em áreas variadas como a psicologia, processamento e compreensão da linguagem (nossa caso), motivação, emoção e comportamento social.

Nos estudos linguísticos, o fenômeno da Causalidade Implícita é explorado para tratar sobre como operam as relações entre informações de ordem gramatical e de ordem discursiva, como podemos visualizar nos trabalhos de Rohde e Kehler (2014) e Rohde (2008). Especialmente nos estudos da Psicolinguística, de acordo com Goikoetxea, Acha e Pascal (2008), é analisado o papel da causalidade na compreensão dos discursos a vista de que as inferências relacionadas a causas são parte do conhecimento em que o falante pode acessar para apreender o sentido dos discursos (Garrod; Terras, 2000; Kintsch, 1988, 1998; McKoon, Greene; Ratcliff, 1993; McKoon; Ratcliff, 1988).

Nesses estudos de caráter experimental, é possível observar a investigação do fenômeno da CI por meio da utilização frequente de tarefas de produção linguística. Essas tarefas consistem em pedir para que o participante complete sentenças (similares às apresentadas anteriormente) com o conector causal “porque” articulando as estruturas. Portanto, a intenção é a de evidenciar uma potencial causa do evento descrito na sentença. Notamos estudos em diversas línguas que partem desse método para investigar o fenômeno da CI, a saber: Goikoetxea, Acha e Pascual (2008) em espanhol; Ferstl, Garnham e Manouilidou (2011) em inglês; Costa, Faria e Kail (2004) em português europeu e Carvalho e Godoy (2021) em português brasileiro. Para além desses estudos, facilmente encontrados nas bases de dados de pesquisa acadêmica, temos também o trabalho de Costa (2003/2004), trabalho pioneiro em investigar a CI em português na sua variante europeia, mas que infelizmente não está disponível em repositórios de teses ou de artigos acadêmicos.

Tais estudos, em geral, possuem o objetivo de formular *corpus* de verbos de Causalidade Implícita tendo em vista a importância do estabelecimento de corpora normatizados a serem utilizados como base para investigações futuras. Especialmente no trabalho de Goikoetxea, Acha e Pascal (2008), os achados demonstram que a CI ocorre de forma consistente nos verbos interpessoais do espanhol. Por exemplo, o direcionamento das tendências verbais por verbo mostrou que a maioria (72 dos 100 verbos em adultos e 66 dos 100 verbos em crianças) apresentaram a direção esperada, além disso, mais da metade (49 dos 72 verbos em adultos e 35 dos 66 verbos em crianças) mostrou uma medida de tendência de moderada a forte (66%–100% no caso de SN1 e 0%–34% no caso de SN2) (Goikoetxea; Pascual; Acha, 2008).

Resultados consistentes sobre a ocorrência da CI também foram apresentados em língua inglesa, no trabalho de Ferstl, Garnham e Manouilidou (2011) os pesquisadores investigaram a CI em 305 verbos do inglês e os achados, de forma geral, revelaram que um grande número de verbos no *corpus* apresenta um viés significativo para SN1 ($n = 127$) ou SN2 ($n = 112$) (Ferstl; Garnham; Manouilidou, 2010).

1.2 Taxonomia dos verbos de causalidade implícita

Os verbos interpessoais podem indicar determinados comportamentos dos participantes envolvidos na ação e/ou estado indicado pelo verbo. No que se refere, particularmente, à categorização dos verbos de Causalidade Implícita, o trabalho de Rudolph (1997) apresenta a *taxonomia de ação-estado revisitada*, originalmente denominada de *revised action-state taxonomy*, a qual divide os verbos que operam com dois participantes envolvidos em um evento em quatro categorias: Agente-Paciente (AP), Agente-Evocador (AE), Estímulo-Experienciador (SE) e Experienciador-Estímulo (ES).

Os verbos AP envolvem um agente – provocador de uma ação de modo voluntário, segundo Rudolph (1997), – e um paciente – ser que passaria por mudanças de estado (e.g., Sofia buscou João porque.../ João buscou Sofia porque...); Os verbos AE implicam um agente praticante de uma determinada ação em resposta ao comportamento ou estado de outrem (e.g., Diego criticou Eloá porque.../ Eloá criticou Diego porque...) enquanto os verbos SE e ES indicam papéis de estímulo (S) – um causador de experiência ou estado mental – e experienciador (E) – alguém que vivencia uma experiência (e.g., SE: Amanda surpreendeu Gabriel porque.../ Gabriel surpreendeu Amanda porque...; ES: Samuel comprehendeu Helena porque.../ Helena comprehendeu Samuel porque...). Desse modo, tais categorias representam as relações semânticas estabelecidas entre os protagonistas do evento descrito pelo verbo, isto é, a divisão dos verbos nessas subclasses tem como base o papel temático ocupado respectivamente pelos sujeitos e objetos das sentenças.

Especialmente, os estudos de Goikoetxea, Acha e Pascual (2008) e Ferstl, Garnham e Manouilidou (2011), ao utilizarem tal taxonomia, hipotetizaram que o comportamento dos verbos de CI poderiam ocorrer da seguinte maneira: os verbos AP e SE se comportariam como verbos com tendência ao SN1, enquanto verbos AE e ES se comportariam como verbos com tendência ao SN2. Os autores reportam efeitos consistentes na interação entre tendências da CI e tipo de verbo em ambos trabalhos: em Goikoetxea, Acha e Pascual (2008), os verbos de estado (SE e ES) apresentaram um seguimento maior as tendências verbais da CI do que os verbos de ação (AP e AE); já em Ferstl, Garnham e Manouilidou (2011), os verbos AE e ES provocaram mais continuações direcionadas ao SN2, os verbos SE geraram mais continuações direcionadas ao SN1 e os verbos AP não resultaram em nenhuma preferência.

Como resultado de tais trabalhos, há hoje um *corpus* estabelecido em espanhol e inglês de verbos com tendências de CI atestadas. Para nos ajudar a construir *corpus* semelhante em português, também adotamos a taxonomia de verbos descrita nesta seção, pois acreditamos que ela fornece uma divisão clara e objetiva das categorias verbais, mostrando-se favorável à manipulação experimental. Além disso, é notável o seguimento dessa categorização em outros trabalhos sobre a CI e a ocorrência de uma possível interação entre os vieses dos verbos de CI com o tipo verbal (Goikoetxea; Pascual; Acha, 2008; Ferstl; Garnham; Manouilidou, 2011). Logo, é pertinente investigarmos se tal interação também ocorre no PB.

1.3 Causalidade Implícita e uma possível interação com gênero

Outro ponto de interesse quando investigamos a CI de um evento é sua interação com o gênero dos participantes da ação descrita (o sujeito e objeto da oração) e o próprio gênero dos participantes da pesquisa. A questão de gênero envolvendo o fenômeno da CI é explorada de forma evidente no trabalho de LaFrance, Brownell e Hahn (1997), em que o objetivo central é investigar se o gênero dos participantes pode afetar a sua percepção como alguém que motiva determinada interação social. A fim de explorar como gênero e linguagem podem estar relacionados, as autoras desenvolvem um levantamento sobre tal questão e direcionam, especialmente, quais são as possíveis tendências envolvidas na relação entre o fenômeno da CI e o gênero. Um apontamento pertinente é em relação aos tipos de verbos utilizados:

Aqui elaboramos estudos em que variamos o gênero de ambos os alvos de interação. Espera-se que o gênero do participante afete as atribuições resultantes, embora vários efeitos sejam concebíveis. Por um lado, o gênero do sujeito da frase pode afetar a causalidade ao interagir com o tipo verbal. Como os estereótipos sugerem que os homens agem enquanto as mulheres reagem (Ruble, 1983), os homens podem estar mais fortemente associados aos verbos de ação e as mulheres aos verbos de estado. Assim, os homens descritos como atores e as mulheres descritas como experienciadoras podem receber mais causalidade do que os pares inversos do sexo alvo e do verbo (LaFrance; Brownell; Hahn, 1997, p.140, tradução nossa)¹.

Notamos, a partir do trecho acima, que pode ocorrer uma interação entre o gênero apresentado pelos participantes do evento com aquilo que é expressado pelo evento (ação ou estado) visto que os estereótipos de gênero podem atuar sobre quem é considerado atuante em um evento e quem é tomado como o experienciador/paciente de um acontecimento. Assim, as autoras destacam que a manipulação do gênero dos elementos de uma sentença e do tipo verbal pode ter efeitos sobre as atribuições causais.

Ademais, outro ponto relevante destacado pelas autoras é de que o gênero dos protagonistas das sentenças experimentais possa afetar as atribuições causais. Segundo LaFrance, Brownell e Hahn (1997), é possível que os homens sejam percebidos como mais causais do que as mulheres independentemente do tipo de verbo uma vez que, em geral, as colaborações dos homens recebem maior atenção e valorização do que as das mulheres (Butler; Geis 1990; Craig; Sherif, 1986; Ridgeway, 1981; Robinson; McArthur, 1982).

O trabalho das autoras objetiva, então, verificar se essas tendências hipotetizadas são realmente apresentadas por meio de uma investigação experimental. Apesar dos resultados de LaFrance, Brownell e Hahn (1997) não corroborarem as hipóteses mais específicas formuladas pelas autoras, foram identificadas interações significativas entre gênero e CI, por exemplo, quando a uma mulher é a receptora da ação dos outros, esta é percebida como quem causou o evento muito mais do que os homens (LaFrance; Brownell; Hahn, 1997). Assim, o trabalho lança questionamentos importantes sobre linguagem e gênero que podem ser aprofundadas e/ou testadas novamente em trabalhos posteriores.

Ao nos debruçarmos sobre trabalhos mais recentes que visam construir corpora normatizados sobre o fenômeno da CI, é notável a seleção do fator de gênero como uma variável de possível interação com o fenômeno linguístico investigado (Goikoetxea; Acha; Pascual, 2008; Ferstl; Garnham; Manouilidou, 2011). Nesses estudos, são considerados como fatores de interação tanto o gênero dos participantes do evento quanto o gênero dos participantes envolvidos na pesquisa.

Em especial, no trabalho de Goikoetxea, Acha e Pascual (2008), que baseou-se em LaFrance, Brownell e Hahn (1997), o gênero de SN1 foi selecionado como uma variável de possível interação com o fenômeno da CI a fim de testar a hipótese de que quando protagonis-

¹ Here we designed studies in which we varied the sex of both interaction targets. Interactant's sex is expected to affect the resulting attributions, although several effects are conceivable. On the one hand, sex of sentence subject might affect causality by interacting with verb type. Because stereotypes suggest that males act while females react (Ruble 1983), males might be associated more strongly with action verbs and females with state verbs. Thus males described as acting and females described as feeling might be assigned more causality than the reverse pair- ings of target sex and verb (LaFrance; Brownell; Hahn, 1997, p.140).

tas masculinos (SN1 – masculino) atuam sobre protagonistas femininos (SN2 – feminino), há uma percepção dos homens como mais causais do que quando as mulheres atuam sobre os homens. Para explorar esse efeito de gênero, as sentenças experimentais formuladas pelos pesquisadores apresentavam o mesmo verbo ora com o sujeito masculino, ora com o sujeito feminino (e.g., “Ana aborreceu Gabriel” e “Gabriel aborreceu Ana”).

Os resultados de Goikoetxea, Acha e Pascual (2008) não apresentaram efeitos significativos para o gênero de SN1. Já no estudo de Ferstl, Garnham e Manouilidou (2011), foram investigadas duas hipóteses em relação ao viés de gênero: 1. Se o gênero dos protagonistas das sentenças afetava o direcionamento causal 2. Se o gênero do participante do experimento poderia modular as atribuições causais. Os resultados do estudo em língua inglesa apontaram os seguintes efeitos de gênero: para a primeira hipótese, houve uma preferência geral por continuações atribuindo a causa ao protagonista masculino das sentenças; para a segunda hipótese, as participantes mulheres apresentaram a tendência de apontar suas continuações para o SN1, independente da ordenação dos SNs, já os participantes homens foram mais propensos a apontarem a atribuição causal ao protagonista masculino das sentenças independente deste ser sujeito (como no caso de verbos AP) ou objeto (como no caso de verbos AE).

Sendo assim, percebemos que há uma divergência entre os estudos nos resultados sobre a interação do gênero do protagonista das sentenças. Além disso, o estudo do espanhol não considerou o gênero do participante como uma variável de interação com o fenômeno da CI, assim, não é possível compararmos as ocorrências entre os idiomas. Por conseguinte, acreditamos ser pertinente adotarmos em nosso estudo as variáveis gênero dos SNs e gênero dos participantes a fim verificar se as tendências relacionadas à interação entre gênero e à Causalidade Implícita se manifestam no PB.

1.4 Estudos de causalidade implícita no PB

Uma vez que vislumbramos alguns estudos de causalidade implícita no inglês e no espanhol, notamos que, no caso do português, há poucos trabalhos detendo-se sobre a temática, especialmente sob um viés experimental. As seções anteriores mostram que a CI é um fenômeno que emerge da interação de diversos fatores: características semânticas dos verbos, gênero dos SNs envolvidos no evento e mesmo gênero dos participantes da pesquisa. Em português, temos estudos que se dedicam ao fenômeno, mas nenhum deles aborda essas variáveis.

No caso do PE, Costa, Faria e Kail (2004) desenvolveram um estudo com dois experimentos – vamos nos deter apenas ao primeiro, que é um experimento de complementação de sentenças como o nosso – para analisar a relação entre informações de viés sintático (estrutural) e semântico (discursivo) na anáfora com verbos de CI.

O experimento 1 foi uma tarefa *offline* de complementação de sentenças. Todas as orações possuíam um verbo de CI em seu núcleo e uma conjunção causal logo em seguida, de forma que cabia ao participante escolher a quem se referir e como fazer isso. Ex.: “No tribunal, a Graça desiludiu o Paulo porque...”. Os autores encontraram que o PE possui duas classes de verbos de CI: com viés SN1 e com viés SN2. As informações semânticas, ou seja, o viés de causalidade implícita, influenciaram o processo anafórico de forma muito mais robusta que os processos sintáticos. É interessante observar como esses processos se dão no PE, pois nos oferecem um vislumbre de como se dá o funcionamento dos processos anafóricos dos verbos de CI no PB.

Nessa mesma linha, o trabalho de Zhang (2019) examina como ocorre o desempenho de falantes de chinês que estão aprendendo Português Europeu como L2 no estabelecimento de anáforas em contexto de causalidade implícita. Assim como no trabalho anterior, o intuito era descobrir como se dá a integração sintaxe-semântica no estabelecimento de correferência. O resultado foi que as informações semânticas – a causalidade – predominaram para os falantes escolherem qual elemento possuía mais probabilidade de ser mencionado, independente de função sintática. A conclusão do autor é que as informações relacionadas à estrutura argumental dos verbos de CI é apreendida muito mais rápido pelos falantes de L2 que as questões estruturais (se eles preferem um pronome pleno ou nulo para retomar o antecedente, por exemplo).

Já no PB, Carvalho e Godoy (2021) realizaram um estudo para construção de um *corpus* com 50 predicados no português brasileiro. As autoras basearam-se nos estudos de Costa (2003/2004) para os estudos de *corpora*, utilizando-se dos mesmos verbos do trabalho em PE com adaptações e acréscimos do PB. No trabalho, foram encontrados 24 verbos com viés de causalidade SN1 e 22 verbos com viés SN2 através de um estudo de complementação de sentenças.

Porém, o estudo se deteve a analisar unicamente o viés de menção para cada predicado considerando sua causalidade implícita, tendo em vista que era um mapeamento inicial. Para o nosso trabalho, procuramos realizar uma continuidade dos estudos de Carvalho e Godoy (2021) considerando o acréscimo de outras análises, as quais já foram explicitadas anteriormente: (i) o papel temático desempenhado por SN1 e SN2 – Agente-Paciente (AP), Agente-Evocador (AE), Estímulo-Experienciador (SE) e Experienciador-Estímulo (ES) -, (ii) o quanto a percepção de causalidade é afetada pelo gênero – masculino e feminino – do sujeito e do objeto. A nossa hipótese, que encontra respaldo nos estudos resenhados nas sessões anteriores, é de que informações semânticas do verbo no que diz respeito a seus papéis temáticos serão decisivas para definir o viés de cada verbo.

A partir deste quadro, realizamos um experimento de continuação de sentenças a fim de constituir *corpus* de CI em PB a ser abertamente disponibilizado à comunidade acadêmica para uso em pesquisas futuras. Diferentemente do que foi feito em estudos anteriores tanto em PB quanto em PE, controlaremos a variável tipo de verbo, seguindo taxonomia de Rudolph (1997). Além disso, a fim de contribuir para a literatura sobre CI para além do português, investigaremos os efeitos do gênero dos SNs e do gênero dos participantes envolvidos na pesquisa. Dessa maneira, também poderemos discutir como o fenômeno da CI é influenciado por fatores culturais.

2 Metodologia

Neste trabalho executamos um experimento de continuação de sentenças com o objetivo de examinar a Causalidade Implícita em 80 verbos do português brasileiro. O experimento de continuação de sentenças consiste em pedir para os participantes completarem frases com primeira continuação que lhes vier à mente. A partir da literatura sobre a CI, se o fenômeno for consistente em PB, esperava-se que uma parte das sentenças direcionassem suas continuações para o sujeito (SN1) e outra parte das sentenças orientem suas continuações para o objeto (SN2). Mais especificamente, esperávamos que verbos AP e SE tendem a atribuir a causa do evento para o sujeito da sentença e verbos AE e ES tendem a atribuir a causa do evento para o objeto da sentença. Com relação ao gênero gramatical de SN1, esperava-se

que os protagonistas masculinos seriam selecionados com maior frequência como causa dos eventos do que protagonistas femininas. Com relação ao gênero dos participantes de pesquisa, havia a expectativa de que homens e mulheres usariam diferentes estratégias na atribuição da causalidade.

2.1 Materiais e Métodos

A fim de testar nossas hipóteses, os estímulos experimentais foram criados manipulando duas variáveis: tipo de verbo e gênero do SN1. A variável tipo de verbo segue a classificação de Rudolph (1997) e apresenta quatro níveis: Agente-Paciente (AP), Agente-Evocador (AE), Estímulo-Experienciador (SE) e Experienciador-Estímulo (ES) (ver Tabela 1). Para cada categoria, foram selecionados 20 verbos, compondo um total de 80 estímulos experimentais. Os verbos escolhidos baseiam-se na seleção para o espanhol feita por Goikoetxea, Acha e Pascual (2008). Inicialmente, selecionamos a tradução de 100 verbos do espanhol usados pelos autores que apresentam cognatos em português. Posteriormente, excluímos 20 verbos do espanhol, pois em alguns casos a tradução não se apresentava adequada para o PB (e.g., *alcanzar*/alcançar; *poner*/colocar; *señalar*/apontar). Ao final, incluímos 29 verbos do português a partir de sinônimos dos verbos em espanhol e mantivemos 51 verbos das traduções do espanhol, assim, totalizando os 80 verbos selecionados para este trabalho. Além disso, para realizar a distribuição dos verbos nas categorias AP, AE, SE e ES, tomamos como base a distribuição feita no estudo do espanhol. Tal estudo em língua espanhola tomou como base a lista de verbos do estudo de Rudolph (1997) e frequência dos verbos em espanhol para realizar essa distribuição (Goikoetxea; Acha; Pascual, 2008).

Para testar nossa hipótese específica sobre a relação entre CI e gênero dos participantes do evento evocado pelo verbo, foram elaboradas duas sentenças experimentais para cada verbo: uma com o SN1 masculino e o SN2 feminino; e outra com o SN1 feminino e o SN2 no masculino. Assim como em Goikoetxea, Acha e Pascual (2008), a oração era seguida pelo conectivo “porque” a fim de explicitar a relação causal presente nas sentenças, portanto, as sentenças seguem a seguinte configuração: “SN1 V SN2 porque...”. Além disso, similar ao trabalho em espanhol, todos os verbos utilizados na estruturação das sentenças foram conjugados no pretérito perfeito. A tabela 1, a seguir, apresenta exemplares das sentenças utilizadas classificadas por tipo de verbo e gênero do SN1.

Tabela 1 – Exemplares de itens experimentais

tipo de verbo	sentença
Agente-Paciente (AP)	SN1-masculino: Leonardo ligou para Fernanda porque... SN1-feminino: Fernanda ligou para Leonardo porque...
Agente-Evocador (AE)	SN1-masculino: Leandro parabenizou Jessica porque... SN1-feminino: Jessica parabenizou Leandro porque...
Estímulo-Experienciador (SE)	SN1-masculino: Anthony incentivou Sophia porque... SN1-feminino: Sophia incentivou Anthony porque...
Experienciador-Estímulo (ES)	SN1-masculino: Davi gostou de Ana porque... SN1-feminino: Ana gostou de Davi porque...

Fonte: Elaborada pelas autoras.

Os nomes próprios selecionados para compor as sentenças são provenientes das listas de nomes mais registrados do Brasil nos anos de 2018, 2019 e 2020 e os nomes mais populares da década de 2000 disponíveis, respectivamente, no *Portal da Transparéncia do Registro Civil* e no *Portal Nomes no Brasil* do IBGE (Instituto Brasileiro de Geografia e Estatística) – dados do censo 2010. Os nomes foram controlados pelo número de sílabas, isto é, buscamos, em cada sentença, apresentar os pares de nomes com o mesmo tamanho.

Os itens experimentais foram divididos em 4 listas na seguinte configuração: cada lista é formada por 10 verbos de cada categoria verbal com 5 de cada gênero (masculino e feminino) em cada categoria.

A execução do experimento se deu na plataforma *Google Forms*. Primeiramente, os participantes foram direcionados para a leitura e aceite do Registro de Consentimento Livre e Esclarecido, no qual apresenta-se informações sobre o experimento e sobre os direitos do participante, conforme orientações do Comitê de Ética em Pesquisa (CEP) da Universidade Federal do Rio Grande do Norte (UFRN)², que aprovou o protocolo de pesquisa. Em seguida, os participantes respondiam a um questionário de dados (idade, cidade de residência, gênero, escolaridade, se falante nativo do PB). Posteriormente a essas etapas referidas, os participantes recebiam as instruções para a execução da tarefa experimental (e.g., “Sua tarefa será completar as sentenças intuitivamente, ou seja, você deverá escrever a primeira continuação que vier em sua mente para completar a frase.”) e executavam um breve treino para se habituar ao experimento. Por fim, os participantes foram direcionados para a tarefa experimental em que cada participante deu continuidade a 40 sentenças apresentadas de forma aleatória variando em gênero e categoria verbal (AP, AE, SE e ES).

2.2 Participantes

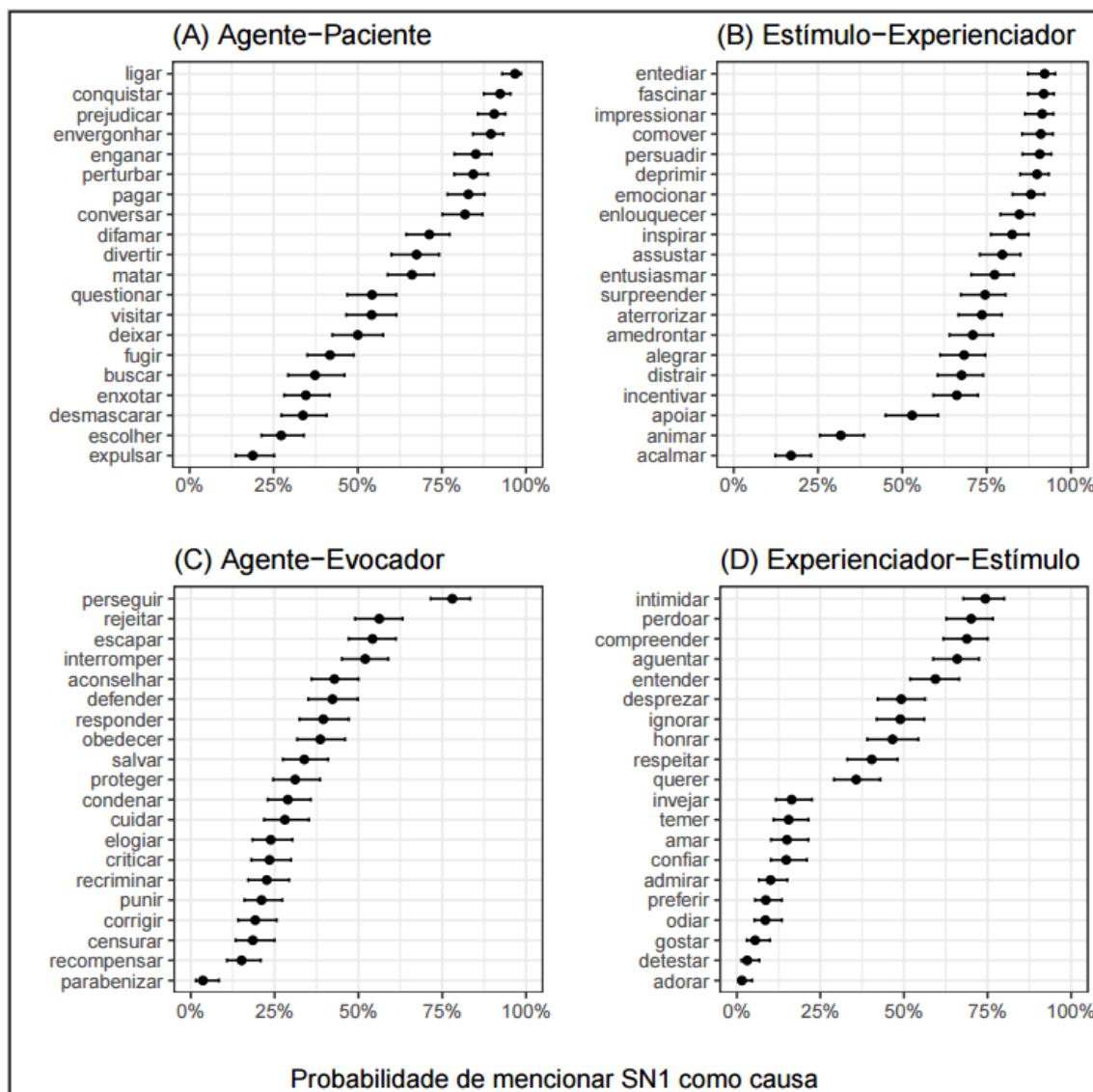
O experimento contou com a participação voluntária de 411 pessoas com a média de idade de 42 anos (variando entre 18 e 66 anos), a escolaridade se enquadra entre ensino fundamental e ensino superior completo. Quanto ao gênero, houve 260 participantes do gênero feminino, 140 do gênero masculino, 5 não-binários e 4 participantes que preferiram não declarar seu gênero. Dados de dois participantes foram descartados por não serem falantes nativos do PB.

3 Resultados

Depois de serem coletados pelo *Google Forms*, os dados foram baixados no *Excel* e organizados e catalogados no *Rstudio*. Os dados foram analisados por anotadores para definir qual o referente foi retomado pelos participantes. Nos casos em que não havia concordância entre os dois, um terceiro fazia a mediação. Quanto aos resultados finais, os dados têm uma distribuição quase igualitária entre respostas do tipo SN1 ($n = 7366$, ou 50.9% dos dados) e SN2 ($n = 7106$, ou 49.1% dos dados). Por outro lado, 12% dos itens são de continuações que não retomam diretamente SN1 e SN2 (encontram-se aqui os casos em que a resposta não revolveu a ambiguidade, retomou outro referente que não SN1 e SN2, retomou ambos referentes ou respostas *nonsense*).

² Número do protocolo da pesquisa: 48214021.3.0000.5537

Figura 1 – Probabilidade de identificar o SN1 como causa do evento de cada um dos verbos

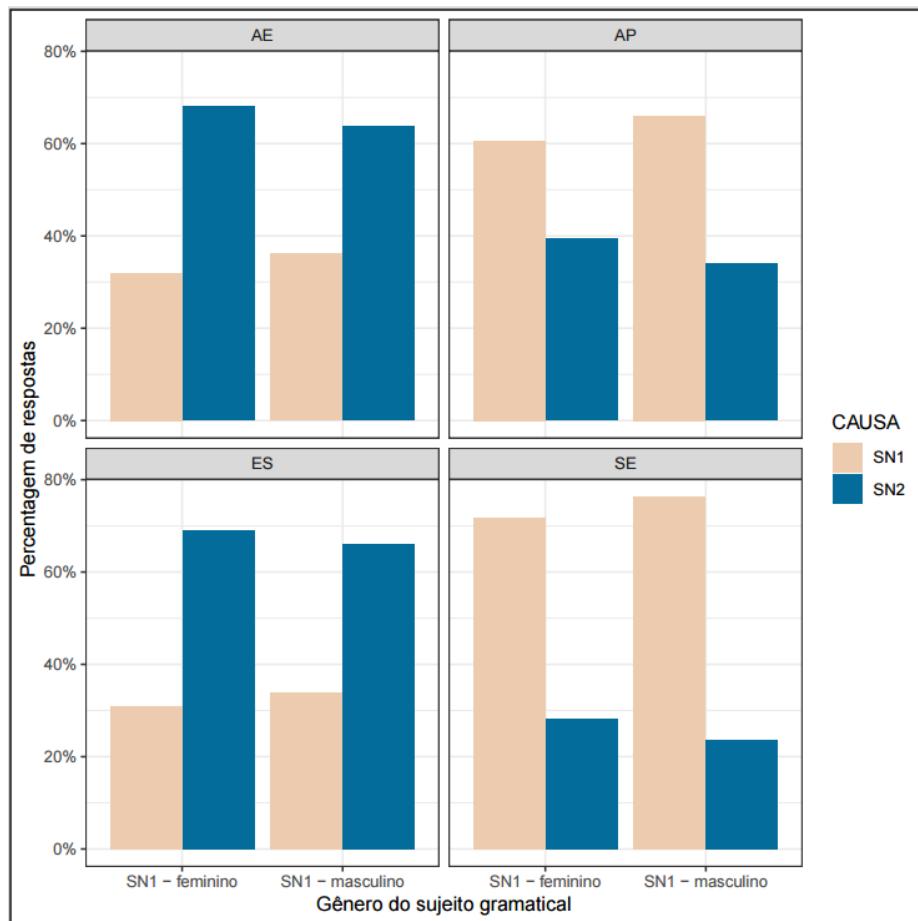


Fonte: Elaborada pelas autoras.

3.1 Taxonomia dos verbos de CI

Além do panorama geral do comportamento dos verbos de causalidade implícita, analisamos também se havia uma influência do tipo de verbo, especificamente, na interpretação causal. Verbos do Tipo Agente-Evocador e Experienciador-Estímulo tiveram um maior número de continuações que apontavam SN2 como causa, enquanto o oposto ocorreu com verbos do tipo Agente-Paciente e Estímulo/Experienciador, como podemos observar na figura 2. Além disso, também procuramos investigar a influência que o gênero de SN1 (sujeito gramatical da oração) poderia ter na percepção da causalidade. Para averiguar se essas tendências realmente se sustentam, realizamos uma análise estatística por meio de modelos mistos (Baayen, Davidson e Bates, 2008). Elegemos esses modelos em vez de modelos lineares menos complexos, como a ANOVA, para incluir em nossa análise os efeitos aleatórios de itens e participantes (cf. Godoy e Nunes, 2020).

Figura 2 – Distribuição de respostas que identificam a causa da oração experimental como SN1 ou SN2 para cada tipo de verbo

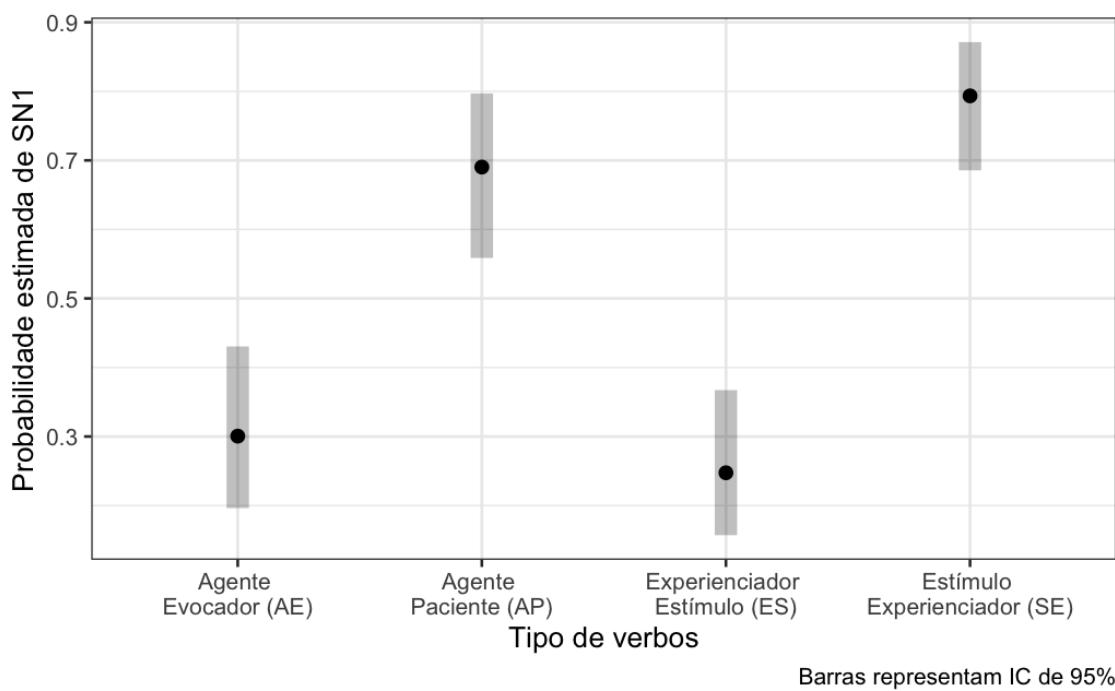


Fonte: Elaborada pelas autoras.

Inicialmente, ajustamos uma regressão logística com a causa como variável resposta, e tipo de verbo, gênero do SN1 e a interação entre esses fatores como variáveis preditoras. Além disso, adicionamos interceptos aleatórios por participantes e itens. Uma comparação por modelos aninhados não identificou efeito significativo da interação ($\chi^2 = 0.975$, $p = 0.8$), mas apontou para um efeitos principais de tipo de verbo ($\chi^2 = 41.04$, $p < 0.0001$) e gênero do SN1 ($\chi^2 = 47.6$, $p < 0.0001$).

Uma análise post-hoc identificou que, com relação ao tipo verbal, não há diferenças significativas entre verbos do tipo AE e SE, e nem entre AP e SE. No entanto, todas as outras comparações se mostraram significativas (cf. Tabela 2). As tendências de cada tipo de verbo evocar SN1 como causa do evento podem ser vistas na Figura 3. Comparações pareadas também demonstraram que há maior probabilidade de indicar o SN1 como agente causador do evento quando ele é do gênero masculino ($b = 0.748$, $p < 0.001$, cf. Figura 4).

Figura 3 – Probabilidade estimada de identificar o SN1 como causa do evento para cada tipo de verbo



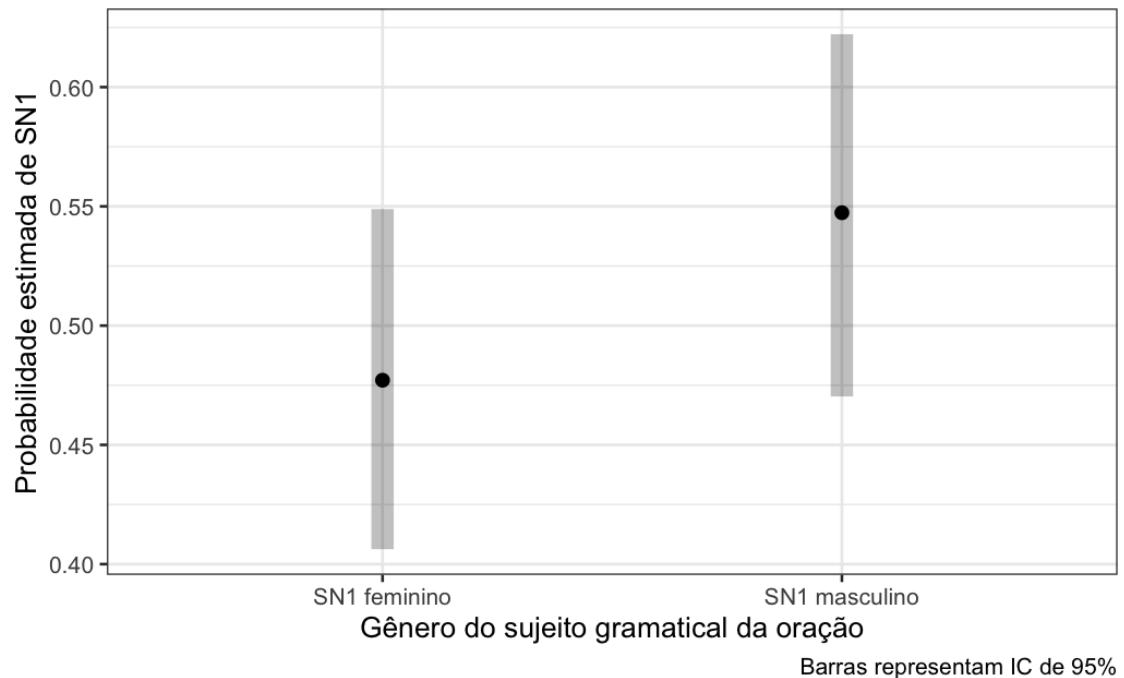
Fonte: Elaborada pelas autoras.

Tabela 2 – Comparações pareadas por tipo de verbo (p-valores ajustados pelo método de Tukey)

comparação	razão de chance	valor z	p-valor
AE/AP	0.19	-4.12	0.00
AE/ES	1.27	0.59	0.93
AE/SE	0.11	-5.40	0.00
AP/ES	6.69	4.69	0.00
AP/SE	0.59	-1.29	0.57
ES/SE	0.09	-5.94	0.00

Fonte: Elaborada pelas autoras.

Figura 4 – Probabilidade estimada de identificar o SN1 como causa do evento para cada tipo de verbo

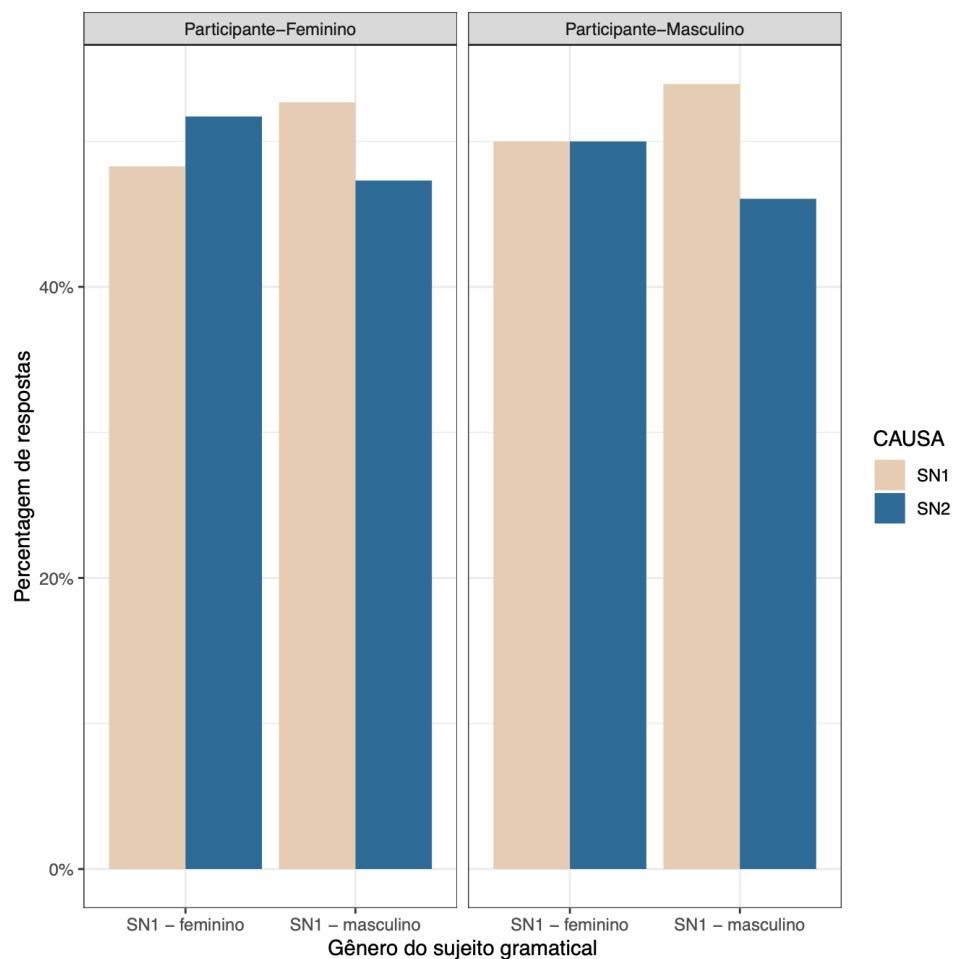


Fonte: Elaborada pelas autoras.

3.2 Relação entre gênero do participante e percepção da causalidade

Procuramos averiguar se há relação entre o gênero do participante e a sua percepção de causalidade. Como o número muito pequeno de participantes não-binários ($n=5$) impede a análise por meio de estatística inferencial, realizamos as análises apenas com os dados dos participantes que disseram se identificar com os gêneros masculino ($n=139$) ou feminino ($n=259$). Não incluímos a variável tipo de verbo nas análises, uma vez que ela não interagiu com gênero do sujeito gramatical em nossas análises anteriores. Nossa análise, portanto, apenas identifica se o gênero do participante teve influência no efeito principal de gênero do SN1 identificado na análise anterior. A Figura 5 mostra a relação gênero do participante x gênero do SN1 na atribuição de causa dos eventos.

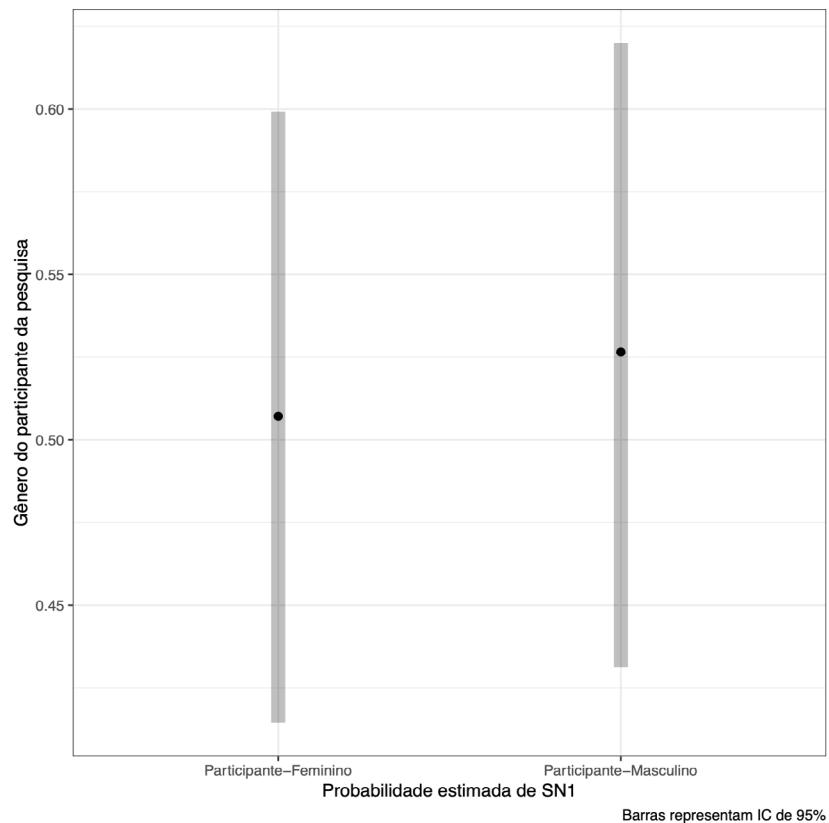
Figura 5 – Distribuição de respostas que identificam a causa da oração experimental como SN1 ou SN2 por gênero do participante



Fonte: Elaborada pelas autoras.

Para testar o efeito das variáveis de interesse, ajustamos novamente uma regressão logística com a causa do efeito (SN1 ou SN2) como variável resposta. As variáveis preditoras foram gênero do SN1, gênero do participante da pesquisa, e a interação entre essas variáveis. Incluímos interceptos aleatórios por participantes e itens. Uma análise por modelos aninhados não apontou efeito significativo da interação entre variáveis ($\chi^2 = 0.037$, $p = 0.847$). Identificamos novamente um efeito significativo do gênero de SN1 ($\chi^2 = 45.2618$, $p < 0.0001$), mas nenhum efeito significativo do gênero do participante da pesquisa ($\chi^2 = 0.986$, $p = 0.321$). A Figura 6 mostra a probabilidade estimada de identificar a causa do evento como SN1 em função do gênero do participante da pesquisa.

Figura 6 – Probabilidade estimada de identificar a causa do evento como SN1 em função do gênero do participante da pesquisa



Fonte: Elaborada pelas autoras.

4 Discussão

Este estudo fornece um *corpus* de verbos de Causalidade Implícita no Português Brasileiro. Com o intuito de evidenciar as causas dos eventos retratados pelos verbos, solicitamos aos participantes a continuarem sentenças estruturadas com o conectivo *porque*. Nossos resultados corroboram descobertas anteriores. Buscamos selecionar verbos semelhantes aos das línguas investigadas anteriormente e mostramos que a maioria dos verbos apresentam uma tendência de continuação de sentenças direcionando a causalidade para o SN1 ou SN2. Este resultado contemplou um grupo de participantes bastante significativo em comparação a estudos anteriores. Tal coleta de dados foi possível devido ao fato de o experimento ser divulgado por meio de listas de e-mail e redes sociais e de ser executado por meio de uma plataforma virtual.

Ao utilizar uma seleção de verbos semelhantes ao estudo do espanhol, podemos observar que nossos resultados são comparáveis aos de Goikoetxea, Pascual e Acha (2008) para verbos em espanhol apesar de utilizarmos formas distintas de aplicação da tarefa experimental. No estudo do espanhol o experimento foi conduzido em uma tarefa de papel e caneta em que o participante continuava sentenças e posteriormente marcava nessas sentenças o SN a qual a continuação se referia. Já em nosso estudo conduzimos a tarefa experimen-

tal por meio de uma plataforma *online* que permite a escrita de respostas longas em que os participantes eram solicitados apenas a continuarem sentenças incompletas, mas a indicação de a quem a continuação fazia referência não foi solicitada.

Em relação à classificação verbal, seguimos a literatura utilizada em estudos semelhantes de CI ao dividir os verbos em quatro tipos: Agente-Paciente (AP), Agente-Evocador (AE), Experienciador-Estímulo (ES) e Estímulo-Experienciador (SE). Notamos que nossos resultados replicaram as descobertas de Goikoetxea, Pascual e Acha (2008) no que diz respeito a tendências gerais para tipo de verbo. Para os verbos AP e SE houve uma clara preferência de direcionar a causa para o SN1, enquanto os verbos AE e ES apresentaram uma notável tendência de direcionamento da causa para o SN2. No entanto, ao observarmos as especificidades de cada classe verbal algumas diferenças se apresentam em relação a estudos anteriores. No estudo de Goikoetxea, Pascual e Acha (2008), os verbos de estado (ES e SE) apresentaram um efeito maior de tendência causal ao comparado com verbos de ação (AP e AE). A classe AP no estudo de Ferstl, Garnham e Manouilidou (2011) não apresentou uma preferência nítida de direcionamento causal. Em nosso estudo, ao comparar as classes verbais de estado e de ação não foi observado efeitos significativos em interação com o fenômeno da CI. Desse modo, observamos diferenças em relação ao comportamento das classes verbais ao interagir com a CI ao compararmos línguas distintas. Tais diferenças manifestadas entre as línguas aqui comparadas demonstram a importância da construção de *corpus* em línguas diversas com o objetivo de verificar quais são as tendências predominantes na investigação do fenômeno da CI.

Este estudo também considerou o gênero como uma possível variável de interação com o fenômeno da CI. No que se refere ao gênero do SN1 das sentenças, nossos achados foram bastante interessantes. Ocorreu uma maior probabilidade de atribuir ao SN1 a causa do evento retratado pelo verbo quando o sujeito das sentenças era masculino independente do tipo de verbo ou do gênero do participante da pesquisa. Tal efeito replica em partes o que foi hipotetizado por LaFrance, Brownell e Hahn (1997): de que homens são frequentemente mais considerados como os causadores de um evento. Essa replicação de efeito para o gênero dos protagonistas da sentença também foi evidenciada no estudo de Ferstl, Garnham e Manouilidou (2011), mas não foi encontrado nos resultados de Goikoetxea, Pascual e Acha (2008). Desse modo, nossos resultados para efeito de gênero do sujeito da oração se assemelham ao da língua inglesa. Já no que se refere à interação do gênero dos participantes, hipotetizamos, conforme LaFrance, Brownell e Hahn (1997), que mulheres e homens poderiam utilizar diferentes estratégias na atribuição causal ao completar as sentenças. No entanto, nenhum efeito significativo em relação ao gênero do participante da pesquisa foi encontrado. Tal resultado diverge com o estudo do inglês de Ferstl, Garnham e Manouilidou (2011) que apresenta efeitos significativos do gênero dos participantes na atribuição causal.

Compreendemos que diferenças entre os contingentes masculinos e femininos das amostras podem ser uma limitação das investigações em considerar o gênero como uma variável visto que tanto no nosso estudo quanto nos reportados em outras línguas há um predomínio de participantes do gênero feminino. No entanto, podemos conjecturar que a consideração, nas análises estáticas, da variabilidade por itens e participantes podem contornar em certa medida essa limitação. O nosso estudo apresenta uma diferença relativamente equilibrada entre os gêneros dos participantes (femininos e masculinos) e uma amostra grande (260 mulheres, 140 homens e 5 não-binários) ao contrário do estudo em espanhol em que há

uma certa disparidade entre os gêneros e uma amostra menor (81 mulheres e 24 homens) e do estudo do inglês que, apesar de uma diferença equilibrada entre gêneros, possui também uma amostra menor (52 mulheres e 44 homens). Dessa maneira, divergências entre os efeitos podem estar relacionadas a amostra de cada estudo visto que diferenças estáveis entre gêneros são melhor retratadas em uma amostra suficientemente grande (Ferstl, Garnham e Manouilidou, 2011). Além disso, podemos especular que as dimensões socioculturais relativas a cada população amostral podem ser fatores de influência ao lidar com questões de gênero, por exemplo, posicionamentos políticos mais conservadores podem levar a atribuições causais mais fortemente associadas com os estereótipos de gênero. Portanto, é interessante que ocorram mais pesquisas para investigar a variável gênero em interação com a CI com a inclusão de diferentes estratégias de respostas nas tarefas experimentais tendo em vista que o fator de gênero se comporta de maneira distinta nas línguas reportadas neste trabalho.

O *corpus* resultante deste trabalho poderá ser utilizado em diversas investigações futuras não só da psicolinguística, mas também de outras áreas como a da psicologia social. Os dados para os verbos de CI do PB podem beneficiar trabalhos que exijam uma diversidade de itens experimentais em que o controle de vieses semânticos tenha de ser levado em consideração. Além disso, o *corpus* é útil em distintas aplicações da psicolinguística, principalmente, em estudos sobre pragmática e relações interpessoais, ao fornecer dados normatizados sobre propriedades semânticas verbais.

Por fim, é notável que o fenômeno da Causalidade Implícita possui algumas questões interessantes em aberto, como é o caso da sua interação com o gênero. Especialmente no PB, o fenômeno também carece de ampliar o alcance de sua investigação, por exemplo, observar questões de compreensão envolvendo a CI. O presente *corpus* poderá, então, facilitar pesquisas futuras envolvendo o fenômeno da Causalidade Implícita.

5 Informações complementares

Os arquivos com os dados brutos, *scripts* de análise e a planilha com informações de CI para cada verbo estão disponíveis no endereço https://osf.io/ypued/?view_only=d90473d918a049f7bbb-756c25e36db60.

Declaração de autoria

Este artigo foi desenvolvido pelas três autoras, membros do Laboratório de Estudos Experimentais em Linguagem (LEELin/UFRN). Todas as autoras participaram do levantamento de dados e colaboraram na redação e revisão do artigo. Especificamente, a primeira autora contribuiu na redação de todas as seções do artigo e na revisão da redação do artigo; a segunda autora contribuiu na redação da revisão teórica, dos resultados e do resumo do artigo; e a terceira autora colaborou no desenvolvimento da análise estatística, na redação dos resultados, orientação e revisão da redação do artigo.

Agradecimentos

Agradecemos à UFRN, ao CNPq e à CAPES, pelas bolsas concedidas para a realização dessa pesquisa, e ao colega de iniciação científica Bruno Cabral, pelo excelente trabalho em conjunto na análise de dados dessa pesquisa.

Referências

- BUTLER, D.; GEIS, F.L. Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social Psychology*, Washington, v. 58, n. 1, p. 48-59, 1990. DOI: <https://doi.org/10.1037/0022-3514.58.1.48>
- CARVALHO, R. S. S.; GODOY, M. Viés de causalidade implícita para 50 predicados do Português Brasileiro. *Calígrama*, Belo Horizonte, v. 26, n. 2, p. 89-105, 2021. DOI: <https://doi.org/10.17851/2238-3824.26.2.89-105>
- COSTA, A.; FARIA, I.; KAIL, M. Semantic and Syntactic Cues' Interaction on Pronoun Resolution in European Portuguese. 5th DISCOURSE ANAPHORA AND ANAPHOR RESOLUTION COLLOQUIUM. Lisboa, Colibri, p. 1-7, 2004. Disponível em: <https://www.academia.edu/21405261/Semantic_and_Syntactic_Cues_Interaction_on_Pronoun_Resolution_in_European_Portuguese_1>. Acesso em: 05 de jun. de 2022.
- CRAIG, J.M.; SHERIF, C.W. The Effectiveness of Men and Women in Problem Solving Groups as a Function of Group Gender Composition. *Sex Roles*, Berlim, v. 14, p. 453-466, 1986. Disponível em: <<https://link.springer.com/article/10.1007/BF00288427>>. Acesso em: 05 de jun. de 2022.
- FERST, EC.; GARNHAM, A.; MANOUILIDOU, C. Implicit causality bias in English: a corpus of 300 verbs. *Behavior Research Methods*, Berlim, v. 43, p.123-135, 2011 DOI: <https://doi.org/10.3758/s13428-010-0023-2>
- GARROD, S.; TERRAS, M. The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language*, San Diego, v. 42, n. 4, p. 526-544, 2000. DOI: <https://doi.org/10.1006/jmla.1999.2694>
- GODOY, M. C.; NUNES, M. A. Uma comparação entre ANOVA e modelos lineares mistos para análise de dados de tempo de resposta. *Revista da ABRALIN*, v. 19, n. 1, p. 1-23, 2020. DOI: [10.25189/rabralin.v19i1.1388](https://doi.org/10.25189/rabralin.v19i1.1388)
- GOIKOETXEA, E.; PASCUAL, G.; ACHA, J. Normative study of the implicit causality of 100 interpersonal verbs in Spanish. *Behavior Research Methods*, Bilbao, v. 40, p. 760-772, 2008. DOI: <https://doi.org/10.3758/BRM.40.3.760>
- IBGE. *Portal nomes no Brasil: Nomes mais populares*. Disponível em:<<https://censo2010.ibge.gov.br/nomes/#/ranking>>. Acesso em: 05 de jun. de 2022.
- KINTSCH, W. The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review*, v. 95, n. 2, p.163-182, 1998. DOI: <https://doi.org/10.1037/0033-295X.95.2.163>
- KINTSCH, W. *Comprehension: A paradigm for cognition*. 1. ed. Cambridge: Cambridge University Press, 1998.

- LAFRANCE, M.; BROWNELL, H.; HAHN, E. Interpersonal Verbs, Gender and Implicit Causality. *Social Psychology Quarterly*, Thousand Oaks, v. 60, n. 2, p. 138-152, 1997. DOI: <https://doi.org/10.2307/2787101>
- MCKOON, G.; GREENE, S. B.; RATCLIFF, R. Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Washington, v. 19, n. 5, p. 1040-1052, 1993. DOI: <https://doi.org/10.1037/0278-7393.19.5.1040>
- MCKOON, G.; RATCLIFF, R. Contextually relevant aspects of meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Washington, v. 14, n. 2, p. 331-343, 1988. DOI: <https://doi.org/10.1037/0278-7393.14.2.331>
- PORTAL DA TRANSPARÊNCIA DO REGISTRO CIVIL. *Portal da Transparência: registro civil*, 2018. Início. Disponível em: <<https://transparencia.registrocivil.org.br/inicio>>. Acesso em: 05 de jun. de 2022.
- RIDGEWAY, C. L. Nonconformity, Competence, and Influence in Groups: A Test of Two Theories. *American Sociological Review*, Amherst, v. 46, n. 3, p. 333-347, 1981. DOI: <https://doi.org/10.2307/2095064>
- ROBINSON, J.; MCARTHUR, L. Z. Impact of Salient Vocal Qualities on Causal Attribution for a Speaker's Behavior. *Journal of Personality and Social Psychology*, v. 43, n. 2, p. 236-247, 1982. DOI: <https://doi.org/10.1037/0022-3514.43.2.236>
- RUDOLPH, Udo. Implicit verb causality: Verbal schemas and covariation information. *Journal of Language and Social Psychology*, v. 16, n. 2, p. 132-158, 1997. DOI: <https://doi.org/10.1177/0261927X970162002>
- VAN DEN HOVEN, E.; FERSTL, EC. The Roles of Implicit Causality and Discourse Context in Pronoun Resolution. *Frontiers Commun*, v. 3, n. 53, 2018 DOI: <https://doi.org/10.3389/fcomm.2018.00053>
- ZHANG, X. *Causalidade implícita e cadeiras correferênciais: produção de frases causais por nativos de chinês aprendentes de PE-L2*. 2019. 143 f. Dissertação (Mestrado em Psicolinguística) – Faculdade de Letras da Universidade de Lisboa. Lisboa, Portugal, 2019.

Guiné Equatorial: política linguística, manutenção e obsolescência das línguas oficiais, étnicas e crioulas em um contexto ibero-africano

Equatorial Guinea: Language Policy, Maintenance and Obsolescence of Official, Ethnic, and Creole Languages in an Ibero-African Context

Gabriel Antunes de Araujo
Universidade de Macau (UM) | Macau | CN
Universidade de São Paulo (USP)
São Paulo | SP | BR
gabriel.antunes@icloud.com
<https://orcid.org/0000-0001-7337-3391>

Ana Lívia Agostinho
Universidade Federal de Santa Catarina (UFSC) | Florianópolis | SC | BR
a.agostinho@ufsc.br
<https://orcid.org/0000-0002-2395-4961>

Resumo: Na República da Guiné Equatorial (RGE), o espanhol, o francês e o português são línguas oficiais. Além disso, línguas africanas (fang, pichi, bube, fa d'Ambô, entre outras) também são usadas pela população. Mais da metade da população do país tem menos de 20 anos e, com o aumento da migração campo-cidade, as línguas dos diferentes grupos étnicos têm ficado restritas a seus territórios tradicionais, ao passo que o fang, o pichi e o espanhol têm expandido seu uso no dia a dia e se tornado as línguas majoritárias no país, graças, sobretudo, à presença nos centros urbanos e à população jovem. Nossa objetivo aqui é descrever e analisar a situação linguística na RGE, destacando o crescimento do pichi e do fang, a sobrevivência ou obsolescência das demais línguas e o papel das línguas oficiais (o espanhol, o francês e o português).

Palavras-chave: Política linguística; línguas oficiais; contexto ibero-africano; pichi; fang.

Abstract: In the Republic of Equatorial Guinea (REG), Spanish, French, and Portuguese are official languages. Additionally, the population also uses African languages (Fang, Pichi, Bube, Fa d'Ambô, among others). With rural-urban migration increasing, languages of different ethnic groups have been restricted to their traditional territories. Alongside Spanish, in recent years, Fang and Pichi have become the most spoken languages, especially in urban settings by an ever-growing young population. Here, our goal is to



describe and analyze the linguistic situation in the country, the maintenance or obsolescence of REG's languages, and the role of the official languages (Spanish, French, and Portuguese).

Keywords: Language policy and planning; official languages; Afro-Iberian context; Pichi; Fang.

1 Introdução

O objetivo deste texto é descrever e analisar as políticas linguísticas da República da Guiné Equatorial (RGE/GE), um país de língua oficial portuguesa e membro da Comunidade dos Países de Língua Portuguesa (CPLP), servindo assim como uma referência sobre a situação linguística do país, uma vez que não há trabalhos com esse intuito na literatura. Os dados sobre as línguas da RGE são fruto de trabalho de campo realizado em Bioko e Ano Bom em 2012 e da pesquisa bibliográfica da literatura disponível.

A RGE está localizada no Golfo da Guiné na África Atlântica. Em um ambiente multilingüístico, o espanhol (esp)¹ e o francês (fra) possuem status de línguas oficiais, a primeira tendo sido oficializada em 1982 e a última em 1989. Adicionalmente, desde 2010, a RGE é também um país de língua oficial portuguesa (por). Portanto, trata-se do único país de colonização 'ibérica' que possui tanto o espanhol como o português como línguas oficiais. Além disso, são faladas por populações do território outras doze línguas: dez línguas bantu – o fang (fan), o bube (bvb), o kombe (nui), o kwasio (nmg), o seki (syi), o bapuku (bnm), o benga (bng), o lengue (bxc), o iyasa (yko) e o gyele (gyi) – e duas línguas crioulas – o pichi (fpe) e o fa d'Ambô (fab).

O texto está organizado da seguinte forma: depois desta *Introdução*, apresentamos um breve resumo da história colonial da RGE e de seu período pós-independência. Em seguida, detalhamos as políticas para as suas dez línguas étnicas, duas línguas crioulas e três oficiais. Posteriormente, apresentamos nossa discussão e análise. As considerações finais concluem o texto.

2 Contexto sócio-histórico

A RGE é formada por um território continental e um conjunto de ilhas. A parte continental, Mbini, anteriormente conhecida como Rio Muni, faz fronteira ao sul e ao leste com a República do Gabão e, ao norte, com a República de Camarões. Os territórios insulares compreendem a ilha meridional de Ano Bom (ou Pagalu), a ilha setentrional de Bioko (anteriormente conhecida como Fernando Pó) e as ilhas Elobey (Grande e Chico/Pequeno) e Corisco, próximas ao

¹ Utilizamos aqui os códigos ISO 639-3 para cada uma das línguas mencionadas. Disponíveis em https://iso639-3.sil.org/code_tables/639/data.

delta do Rio Muni. Essas ilhas oceânicas, Bioko e Ano Bom, são territorialmente descontínuas e, entre elas, está o território da República Democrática de São Tomé e Príncipe (ver Imagem 1). A ilha de Bioko concentra mais de um quinto da população total do país, de 1,89 milhão (Inege, 2021; Worldbank, 2024 – estimativas).

Imagem 1 – Território da República da Guiné Equatorial (parte continental e ilhas de Bioko e Ano Bom).



Fonte: Cia (2021).

Nominalmente, até 1778, o território da GE pertencia ao Império Português. O início da colonização europeia na região da RGE encontra populações diversas no continente e na Ilha de Bioko. No continente, os portugueses se limitaram à construção de entrepostos para o comércio de escravizados, porém em desvantagem com os já estabelecidos portos de Daomé (ao norte) e de Angola (ao sul). A ilha de Bioko, no período colonial denominada Fernão do Pó ou Fernando Pó, era habitada, na chegada dos colonizadores europeus, pelo povo bube. Todavia, por ser considerada insalubre pelos europeus, a ilha foi tratada como um território menos importante, tanto pelos portugueses, como pelos espanhóis, que assumiram sua soberania *de iure* em 1778. Anteriormente, os portugueses tentaram implantar engenhos de cana-de-açúcar, com pouco sucesso e já sem vestígios de sua presença, no momento da chegada dos espanhóis. Porém, a ilha foi utilizada como entreposto de escravizados, tanto por portugueses, como por holandeses, empurrando a população bube para o seu interior.

(Burton, 1992; Sundiata, 2003). O Tratado de El Pardo² (1778, Espanha) pacificou as relações entre o Império Espanhol e o Português, ao finalizar a negociação de territórios e hostilidades em vários pontos da fronteira luso-espanhola *de facto* da América do Sul. De acordo com o Tratado, Portugal cedia à Espanha as ilhas de Fernão do Pó e Ano Bom, bem como o território continental litoral da atual RGE (ou a Costa de los Esclavos), em troca da paz em sua fronteira meridional. Assim, Portugal obtém territórios na América do Sul, consolidando as futuras fronteiras do Brasil. Para os espanhóis, o controle de Fernando Pó, próximo ao Delta do Níger, na atual Nigéria, daria acesso ao mercado de escravos. No entanto, o domínio espanhol na região, após a assinatura do Tratado, foi pouco efetivo, ao ponto de, em 1822, os espanhóis arrendarem a ilha de Fernando Pó ao Império Britânico (Sundiata, 1974). Em 1855, a Espanha retoma a ilha com o fim do *leasing*, e envia ao território, em 1858, o seu primeiro governador, Carlos Chacón (Álvarez-Chillida; Nerín, 2018; Garcia Cantús, 2004; Nerín, 2010a). Mas, sem uma política econômica ou social efetiva, que só surgiria, via missionários, no final do século, a colonização das ilhas é vacilante e a da costa continental, inexistente. Contudo, a inabitada ilha de Ano Bom, descoberta entre 1493 e 1501 e colonizada a partir de 1503, é local de uma

² Tratado de El Pardo (1778, ESPAÑA, s.d.), artigo 13, em português: “Dezejando S. S. M. M. Fidelissima e Catholica promover as Ventagens do Commercio dos Seus respectivos Vassalos, os quaes podem verificar-se no que reciprocamente fizeram de compra, e venda de Negros, sem ligar-se a Contratos, e Assentos prejudiciaes, como os que em outro tempo Se fizeram com as Companhias Portugueza, Franceza, e Ingleza, as quaes foi preciso extinguir, ou anular: Convieram os dous Altos Príncipes Contratantes, em que para lograr aqueles, e outros fins, e compensar de algum modo as Cessões, restituições, e renuncias feitas pela Corona de Espanha no Tratado Preliminar de Limites de 1º de Outubro de 1777, cederia S. M. Fidelissima, como de facto tem cedido, e cede por Si, e em Nome de Seus Herdeiros, e Successores à S. M. Catholica, e aos Seos Herdeiros, e Successores na Corona de Espanha, a Ilha de Anno Bom na Costa d’Africa, com todos os Direitos, Possessões, e Acções, que tem à mesma Ilha, para que desde logo pertença aos Dominios Espanhoes, do proprio modo que até agora tem pertencido aos da Corona de Portugal: E assim mesmo todo o Direito, e acção, que tem, ou pode ter a Ilha de Fernando do Pó, no Golfo da Guiné, para que os Vassalos da Corona de Espanha Se possão estabelecer nella, e negociar nos Portos, e Costas opostas à dita Ilha, como São os Portos do Rio Gabão, dos Camarões, de São Domingos, Cabo Formozo, e outros daquele Distrito; Sem que por isso Se embarece, ou estorve o Commercio dos Vassalos de Portugal, particularmente dos das Ilhas do Príncipe, e de Santo Thomé, que ao prezente vão, e que no futuro forem a negociar na dita Costa, e Portos, comportando-se nelles os Vassalos Portuguezes, e Espanhoes, com a mais perfeita harmonia, Sem que por algum motivo, ou pretexto Se prejudiquem, ou estorvem huns aos outros”.

Em espanhol: “Deseando Sus Majestades Católica y Fidelísima promover las ventajas del comercio de sus respectivos Súbditos, las cuales pueden verificarse en el que reciprocamente hicieren de compra, y venta de Negros, sin ligarse á Contratas, y Asientos perjudiciales, como los que en otro tiempo se hicieron con las Compañías Portuguesa, Francesa, y Inglesa, los quales fué preciso cortar, ó anular, se han convenido los dos Altos Príncipes Contrayentes en que para lograr aquellos, y otros fines, y compensar de algun modo las Cesiones, Restituciones, y renuncias hechas por la Corona de España en el Tratado Preliminar de Limites de 1º de Octubre del 1777. cederia S. M. Fida, como de hecho ha cedido, y cede por si, y en nombre de Sus herederos y sucesores a S. M. Católica, y los suyos en la Corona de España la Isla de Anno bon, en la costa de Africa, con todos los Derechos, posesiones, y acciones, que tiene a la misma Isla, para que desde luego pertenesca a los Dominios Españoles del propio modo que hasta ahora ha pertenecido á los de la Corona de Portugal; y así mismo todo el derecho, y acción que tienes, ó puedes tener á la Isla de Fernando del Pó en el Golfo de Guinea, para que los Vassallos de la Corona de España se puedan establecer en ella, y negociar en los Puertos, y Costas opuestas á la dicha Isla, como son los Puertos del Rio Gabaon, de los Camarones, de Sto Domingo, de Cabo Fermozo, y otros de aquel Distrito, sin que por eso se impida ó estorve el Comercio de los Vassallos de Portugal, particularmente de los de las Islas del Príncipe, y de Sto Tomé, que al presente van, y que en lo futuro fueren á negociar en dicha Costa, y Puertos, comportandose en ellos los Vassallos Españoles, y Portuguezes con la mas perfecta armonia, sin que por algun motivo, ó pretexto se perjudiquen, ó estorven unos á otros.” (España, 1778, art. 13).

experiência colonizatória portuguesa, do ponto de vista agroindustrial, fracassada (Caldeira, 2006). A partir da ilha de São Tomé, escravizados africanos foram levados para a pequena ilha de Ano Bom (de apenas 17,5 km²). No entanto, uma indústria da cana-de-açúcar se mostrou inviável no território, devido a uma série de fatores, entre eles a pouca produtividade, a concorrência com São Tomé e a distância das rotas marítimas. Por fim, com a resistência das populações africanas transplantadas, os habitantes da ilha se isolaram do mundo. A língua levada de São Tomé acabou por se tornar o *fa d'Ambô* (fab). Tanto em Ano Bom, como em Fernando Pó, os colonizadores europeus encontraram feroz resistência dos povos locais, os anobonenses e os bube, respectivamente (Das Neves, 1991). Nesse sentido, Crespi (2010, p. 9) declara que os territórios portugueses repassados à Coroa Espanhola não seriam de fácil colonização, pois 'quedaba en evidencia que Portugal estaba cediendo la posesión de un territorio poco apto para la colonización y habitado por naturales hostiles'³. No continente, a colonização espanhola enfrentaria maiores dificuldades com as políticas dos reinos locais e a impossibilidade de extrair escravizados em grande quantidade, graças aos bloqueios britânicos no Atlântico no século XIX. Assim, somente no biênio 1926-27 que se inicia a colonização da parte continental da GE, após avanços tímidos nas ilhas de Corisco, Elobey Grande e Elobey Chico e no continente, na foz do rio Muni, com o povo benga, em 1843 (Nerín, 2015). Ao longo do século XX, a Espanha consegue desenvolver atividades econômicas de extração de madeira, e produção de cacau e café, principalmente em Fernando Pó, com mão-de-obra assalariada, importada sobretudo do continente. A população de origem europeia, em 1960, não passava dos 3% (Álvarez-Chillida; Nerín, 2018). Os altos custos de produção e a baixa competitividade das atividades econômicas da Colônia faziam com que seus produtos tivessem como destino quase que exclusivamente a Espanha, obrigada a consumir as exportações da Guiné. Ao mesmo tempo, o Governo Central investia na melhoria das condições de vida da população com obras públicas e investimento em educação e saúde. A independência política do território em 1967 resultou na cessão desses investimentos e um forte golpe no desenvolvimento da região, ainda segundo (Álvarez-Chillida; Nerín, 2018).

O modelo de colonização implantado pela ditadura franquista (1936-1975) desde os anos 1930 na GE dividia a população entre três grupos: os espanhóis, os 'emancipados' e os nativos ou indígenas. Aos primeiros eram dados todos os direitos; ao grupo dos emancipados, alguns, mas como cidadãos considerados de segunda-classe. Os nativos, por sua vez, não eram vistos como cidadãos de direito, ficando na base da pirâmide social. Os dois primeiros grupos formavam uma minoria, ao passo que os indígenas eram a maioria, com quase 90% da população. Processos de aculturação, geralmente conduzido pelas missões religiosas (católicas), eram responsáveis por 'emancipar' os indígenas, ou seja, 'civilizá-los'. Poucos

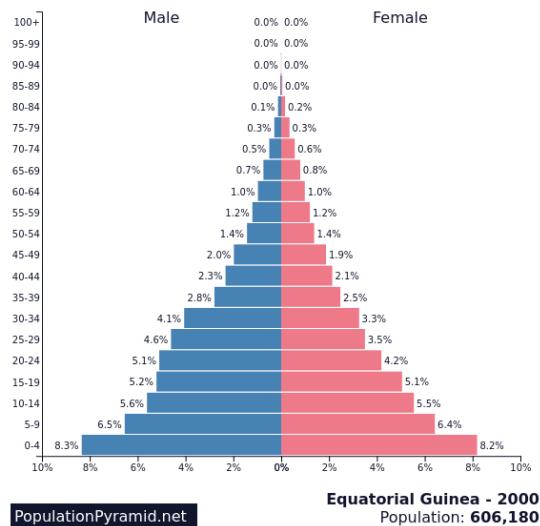
³ Crespi (2010, p. 13, nota de rodapé 19) vai mais além e sugere que o Império Português tenha cometido um estelionato contra o Reino de Espanha: 'Las autoridades portuguesas no brindaron información fidedigna sobre la ubicación y situación de las islas [Ano Bom e Fernando Pó] que decían donar. En la Instrucción Reservada del Tratado del Pardo, por ejemplo, se sitúa inexactamente la latitud de las islas. Además, suponía que Annobón era mayor que Fernando Poo, cuando era a la inversa: 17 km cuadrados contra los 2017 de la segunda. Annobón, que estaba a 400 km de la costa de Gabón, no tenía tierras cultivables, su población era diserta y las corrientes marinas que la rodeaban eran con derrota al Brasil. Difícilmente fuera conveniente para el comercio español. Fernando Poo, si bien estaba cerca de la costa no tenía comunicación directa con la otra isla, siendo forzoso pasar primero por Príncipe y Santo Tomé. Poblada por la etnia bube, se mantuvo durante siglos sin presencia efectiva europea por lo que su colonización se complicaba aún más.'

equato-guineenses tinham acesso ao ensino médio e superior e não foi formada uma classe média com educação técnica, antes da independência, com grande prejuízo para o futuro da jovem nação (Álvarez-Chillida; Nerín, 2018, p. 21-22).

Ademais, a pequena elite que conduziu o processo de independência, cresceu com a cultura fascista do regime franquista, distante dos ideais civilizatórios dos regimes democráticos. Dessa forma, as duas ditaduras pós-independência na Guiné Equatorial, a de Francisco Macías Nguema (1969-1979) e a de seu sobrinho e algoz, Teodoro Obiang Nguema Mbasogo (1979-época atual), são frutos do modelo franquista, com elementos sincréticos das cartilhas dos ditadores africanos dos últimos cinquenta anos. Macías aterrorizou a Guiné Equatorial com expulsão de estrangeiros, genocídio de opositores (um quinto da população pode ter sido assassinada). Um terço da população emigrou a partir de uma campanha ostensiva de destruição da economia do cacau, banimento das religiões, perseguição a intelectuais e opositores, além de proibições como o uso de sapatos ou acesso a escolas e hospitais (Kenyon, 2018). Obiang, por sua vez, governador da Ilha de Bioko, chefe da polícia política e líder do extermínio de opositores sob Macías, mesmo com a riqueza oriunda da exploração do petróleo, após se tornar presidente perpétuo, mantém a população da RGE em estado de miséria, com péssimos indicadores sociais, apesar da mais alta renda per capita (em 2022, US\$ 8,052) da África (Worldbank, 2024).

De acordo com os dados do Banco Mundial, a expectativa de vida (em 2021) era de 61 anos, metade da população não tinha acesso a água tratada, quatro de cada dez crianças em idade escolar estavam fora da escola, um terço da população não possuía acesso à eletricidade, somente 26% tinham acesso à internet, 80% do orçamento era gasto em obras de infraestrutura (estradas, aeroportos etc.) e menos de 3% era investido em educação e saúde juntos, quase 5% da população convivia com o vírus que causa AIDS (12º em 107 países com dados divulgados), dois terços da população vivia abaixo da linha da pobreza (Worldbank, 2024) . Ademais, os dados disponíveis (de 2010) sobre alfabetização indicavam que 97% da população com mais de 15 anos sabia ler e escrever (Worldbank, 2024). Cerca de 26% da população vive na zona rural, contudo 71% da população declara viver da agricultura. Segundo o Censo de 2015, 17,1% da população é estrangeira (Inege, 2021, p. 10). A indústria do petróleo e gás e a construção de infraestrutura tem atraído a maior parte dos estrangeiros. Milhares de jovens do país estão desempregados, e não possuem as qualificações necessárias e sequer têm acesso às informações sobre os empregos. Ao mesmo tempo, devido ao baixo investimento em educação, principalmente técnica e vocacional e à ausência de recursos humanos na área de educação e de administração (gerentes juniores e seniores), o país é obrigado a recorrer à mão-de-obra estrangeira. A pirâmide populacional da RGE, no ano 2000, Imagem 2, representava uma típica situação de um país em vias de desenvolvimento na África: grande população, 50,8% do total, abaixo de 20 anos de idade (De Wulf, 2024).

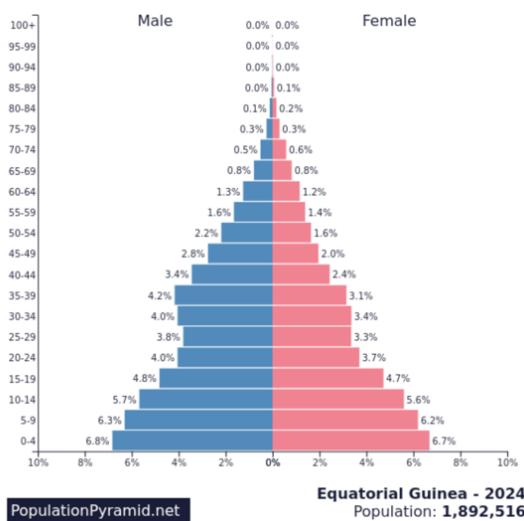
Imagen 2 – Pirâmide populacional da RGE, ano 2000



Fonte: De Wulf (2024).

Já a pirâmide de 2024, Imagem 3, indica um grande número de homens na faixa etária de 20 a 44 anos, idade típica dos trabalhadores estrangeiros que deformam a pirâmide (De Wulf, 2024).

Imagen 3 – Pirâmide populacional da RGE, ano 2024



Fonte: De Wulf (2024)

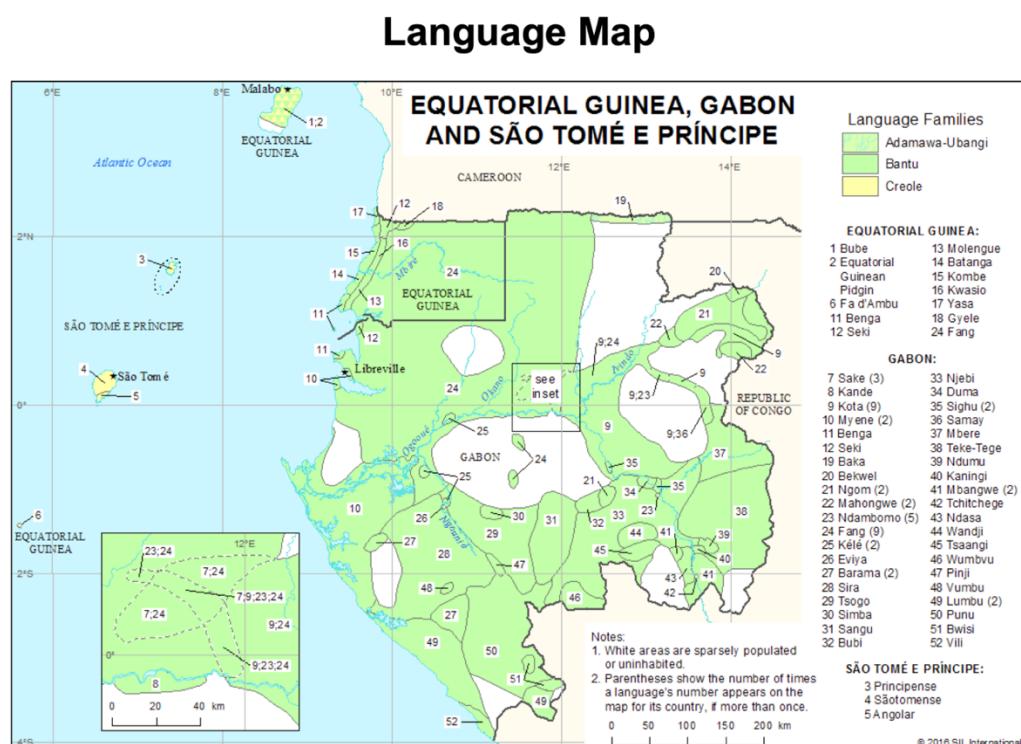
Dessa forma, a história colonial da RGE, associada a seu baixo investimento em educação, terá impacto nas políticas linguísticas do país. Ao mesmo tempo, migração e aumento da população urbana, posto que em 1990, 35% da população vivia nas cidades, ao passo que em 2022, 74% da população de um total de 1,89 milhão de habitantes (cf. Worldbank (2024)) vive nas aglomerações urbanas, ou pelo menos, uma maior migração campo-cidade, também afetam a

transmissão das línguas étnicas (faladas principalmente nas zonas rurais) e os usos linguísticos nas cidades, privilegiando o espanhol e o bubi. Retornaremos a esses pontos na próxima seção.

3 Línguas da Guiné Equatorial

Na RGE, há três grupos distintos de línguas: as línguas étnicas, as línguas crioulas e as línguas oficiais (europeias). No primeiro grupo, temos as línguas faladas por populações que ocupavam, antes da colonização europeia, os territórios de Bioko, o continente e as ilhas próximas à foz do Rio Muni. A ilha de Ano Bom era, antes da chegada dos portugueses, desabitada. As línguas crioulas são faladas principalmente nas ilhas de Bioko e Ano Bom (ver Imagem 4). Por sua vez, na RGE há três línguas europeias oficiais: o espanhol (a língua oficial de facto), o francês e o português⁴. As duas últimas foram oficializadas em 1989 e 2010, respectivamente, embora nenhuma delas apresente uso efetivo em contextos oficiais.

Imagen 4 – Línguas da RGE, segundo o Ethnologue



Fonte: Eberhard; Simons; Fennig, 2022.

⁴ A oficialização do português é mencionada no documento Proyecto de Ley Constitucional (Gobierno-De-Guinea-Ecuatorial, 2010), contudo a Constituição da RGE (Ley Fundamental de Guinea Ecuatorial), cuja última versão foi promulgada em 16 de fevereiro de 2012, em seu artigo 4º trata das línguas oficiais, não menciona explicitamente o português: 'Las lenguas oficiales de la República de Guinea Ecuatorial son en Español, el Francés y las que la Ley determine. Se reconocen las lenguas autóctonas como integrantes de la cultura nacional.' (Gobierno-De-Guinea-Ecuatorial, 2012, p. 4).

Neste trabalho, utilizamos a tipologia de Crystal (2002, p. 19-20)⁵ na qual é necessário considerar alguns fatores indicativos do nível de ameaça linguística, como (i) o número de crianças que adquirem a língua; (ii) as atitudes da comunidade face à língua; e (iii) o grau de impacto de outras línguas na comunidade. Assim, há três níveis de ameaça, fazendo com que as línguas possam ser classificadas como: em segurança, ameaçadas ou extintas. É relevante também considerarmos a nomenclatura de Krauss (1992, p. 4) para as línguas moribundas, que seriam aquelas que não são mais adquiridas pelas crianças como língua materna, ou seja, não há mais a transmissão intergeracional natural. Adicionalmente, é preciso considerar o nível de impacto, por parte de outras línguas, como o espanhol, o fang e o pichi, em diferentes esferas sociais.

3.1 Línguas étnicas

As línguas étnicas são faladas pela população do território e de seu entorno desde antes da chegada dos colonizadores europeus⁶. Com exceção do fang (fan) – língua com cerca de 650 mil falantes nativos –, essas línguas são empregadas por populações relativamente pequenas, variando entre menos de mil e no máximo cem mil falantes. Muitas vezes, são faladas em uma vasta região, porém têm um estatuto minoritário no território, devido à artificialidade das fronteiras na África subsaariana promovidas pela Conferência de Berlim –1884-1885 (Craven, 2015), o que gerou a separação de comunidades falantes de uma mesma língua. Contudo, no caso da RGE, algumas de suas línguas étnicas são limitadas ao território, com exceção do fang, do kombe, do fa d'Ambô e das línguas de origem indo-europeia.

3.1.1 Fang

O fang (fan) é a língua bantu (subgrupo A70) considerada como elemento de unidade nacional, sendo falada por cerca de 80% da população da RGE. A língua é também conhecida como pangwe e possui vários dialetos (fang do sudeste, ntoumou-fang, okak-fang, mekê-fang, mvaïe-fang, atsi-fang, nzaman-fang, nveni-fang). Segundo Eberhard; Simons; Fennig (2022), o fang possui aproximadamente um milhão de falantes L1 e cerca de 35 mil usuários L2. Além da RGE, a língua também é falada no Gabão, nos Camarões e na República do Congo. Não havia falantes de fang nas ilhas de Bioko e Ano Bom até 1930. Contudo, com a migração forçada dos fang para Bioko nos anos de 1930, a situação se alterou um pouco (Bibang Oyee, 1990; Nerín, 2010b). Os dois ditadores, Macías e Obiang, pertencem ao grupo étnico fang (da região continental). No regime de Macías e nos primeiros anos de Obiang, o fang foi, segundo

⁵ Segundo Crystal (2002), os principais fatores indicativos para a classificação de uma língua em processo de extinção são o número de crianças que adquirem a língua como materna, a atitude da comunidade face à língua em questão e, por fim, o grau de impacto de línguas maioritárias na comunidade linguística em questão. O autor define níveis de classificação para as línguas: *segura*, *ameaçada* e *extinta*. Para Crystal (2002), a categoria adicional língua *moribunda* para a língua que não está sendo mais aprendida como língua materna pelas crianças de Krauss (1992, p. 4) captura uma noção que vai além de um mero estágio de ameaça, porque tal classificação, baseada em uma analogia com as espécies de animais que são incapazes de se reproduzir, aborda a característica principal de línguas dessa categoria: a inviabilidade de uma transmissão intergeracional (cf. Agostinho; Bandeira; Araujo, 2016).

⁶ Utilizamos aqui como fonte-base os dados do Ethnologue (Eberhard; Simons; Fennig, 2022), do Glotolog (Hammarström; Forkel; Haspelmath; Bank, 2020) e da classificação das línguas da Guiné Equatorial de Guthrie (Maho, 2009) e Beban; Atindogbe; Domche; Bot (2007).

Nerín (2010b) a língua oficial do país até 1982⁷. Dessa forma, a língua também se beneficia por ser a língua do grupo dominante, o que aumenta seu prestígio.

A primeira descrição do fang foi publicada no final do século XIX por Bates (1899). Bibang Oyee (1990) oferece também uma gramática (científica) do fang, embora o título de seu trabalho (*Curso de lengua fang*) sugira uma gramática pedagógica. Além desses trabalhos, o fang também conta com trabalhos lexicográficos, como Ella (2007), que contém uma lista de palavras fang/francês/inglês e o importante dicionário de Bibang Oyee (2014).

Além de ser uma língua amplamente usada no território, o fang é veiculado nos ambientes familiares e coloquiais, e igualmente empregado em algumas transmissões de rádio. Ao mesmo tempo, é amplamente empregado na literatura oral, no folclore e na música (laica e religiosa), e nos espaços urbanos como *lingua franca*. Nos últimos anos, os Centros Culturales de España da AECID (*Agencia Española de Cooperación Internacional para el Desarrollo*) oferece cursos de língua fang e apoia atividades multimídia dos artistas locais, bem como uma série de atividades inclusivas das comunidades étnicas da RGE, como observado em trabalho de campo. Não há políticas oficiais de promoção do Governo da RGE para a língua fang.

3.1.2 Bube

A língua bube ou bubi (bvb), böbë (autodenominação), também conhecida como bobe, boobe, bohobé, boombe, adeeyah, edeeyah, adija, ediya e fernandiano é uma língua bantu (subgrupo A31), sendo falada por cerca de cinquenta mil pessoas na RGE, principalmente na ilha de Bioko (Eberhard; Simons; Fennig, 2022). Os bube imigraram do continente para a ilha de Bioko há cerca de dois mil anos⁸ (Gelabert; Ferrando-Bernal; De-Dios; Mattorre *et al.*, 2019). Há uma extensa literatura linguística, que se inicia ainda no século XIX⁹ sobre o povo e a língua bube, como por exemplo, Clarke (1841; 1846; 1848), Juanola (1890), Abad (1928), Aymemi (1928), Rurangwa (1987), os trabalhos de Boleka (1986; 1987; 1991; 2009), entre outros, além de uma série de estudos comparativos pan-bantu nos últimos anos. Segundo Eberhard, Simons e Fennig (2022), a língua ainda possui alguns falantes monolíngues, porém a maioria usa o pichi e o espanhol (sendo que 80% são alfabetizados nessa língua), todos os adultos usam a língua, embora muitos jovens prefiram o pichi e o espanhol, o que coloca um futuro desafiador para o bube. Não há políticas oficiais de promoção do Governo da RGE para a língua bube.

⁷ Gomashie (2019, p. 3) corrobora essa afirmação ao constatar que: "From the period of 1968 to 1979, the use of Spanish was prohibited (...) in favour of Fang but it was still the language of international relations and the judiciary, as it was the only written language while the local languages were of oral tradition." Adicionalmente, Nistal Rosique (2007, p. 73-74) afirma que o espanhol chegou a ser proibido por Macías, sendo retomado, como língua de instrução em 1979: "En el periodo de once años que va desde 1968 a 1979, se produjo un enorme daño, no sólo a la lengua española sino a todo el sistema educativo nacional. Fue un periodo sanguinario en el que la educación y la cultura en general, y la lengua española en particular, experimentaron un enorme retroceso." Ademais, somente a Constituição de 1982 viria a oficializar o espanhol como língua do Estado. Igualmente, com o Governo Obiang (1979 até hoje), as escolas voltam a receber alunos, pois o presidente Macías havia banido a educação formal.

⁸ Gelabert *et al.* (2019, p. 1-2) afirmam, citando os estudos de Molino, que 'it is currently accepted that the Bube agriculturalists arrived from the mainland in dugout canoes about 2000 years ago during the Late Neolithic'. Sobre a expansão bantu, a partir da região continental do Golfo da Guiné, ver Koile *et al.* (2022).

⁹ No século XIX, os trabalhos foram publicados principalmente por missionários, um padrão regular, como nas demais línguas descritas na RGE neste período. Ver o estudo de Castillo-Rodríguez (2015) sobre a linguística missionária em Fernando Pó no século XIX.

3.1.3 Kombe

A língua kombe (nui), também grafada combe, ou ngumbi é uma língua bantu (subgrupo A33b) minoritária, sendo falada por cerca de dezesseis mil pessoas na RGE, principalmente na ilha de Corisco e nas regiões adjacentes da costa, e mais alguns milhares na República dos Camarões. Há relatos segundo os quais a língua ndowe ou ndowé seria um dialeto do kombe (Ecuatorial, 2018, p. 9-10). Embora a língua seja falada por uma população relativamente pequena, possui grande vitalidade e é usada amplamente pela comunidade que, como é o caso comum na RGE, também domina outras línguas como o espanhol, o fang e o pichi (Eberhard; Simons; Fennig, 2022). Além de publicações religiosas como o *Novo Testamento*, há poucos estudos científicos publicados sobre essa língua, dos quais se destacam o dicionário do padre Fernandez (1951), o esboço gramatical publicado no dicionário de Fernandez por Maguga (1951), e os trabalhos de Ikuga Ebombebombe (1973a; 1973b), Elimelech (1976) e Ogouamba (1992). Não há políticas oficiais de promoção do Governo da RGE para a língua kombe.

3.1.4 Kwasio

A língua kwasio (nmg), também conhecida como bujeba, mabea, mabi, magbea, mgoumba, mvumbo, ngumba, ngoumba, ngumba é uma língua bantu (subgrupo A81) minoritária, sendo falada por cerca de quinze mil pessoas na RGE, principalmente no norte da parte continental. Embora relacionadas, trataremos o kwasio como uma língua distinta do gyele -kwola (gyi), falada nas florestas interiores da República do Camarões, pois são consideradas línguas distintas (Eberhard; Simons; Fennig, 2022). Embora a língua kwasio seja falada por uma população relativamente pequena, possui grande vitalidade e é usada amplamente pela comunidade que também domina outras línguas como o espanhol, o fang e o pichi (Eberhard; Simons; Fennig, 2022). Além de publicações religiosas como o *Novo Testamento*, a língua kwasio possui alguns textos de descrição linguística, dos quais se destacam Skolaster (1910), González Echegaray (1960), Massaga (1971), Lemb (1974), Dieu (1976) e Um (2002). Não há políticas oficiais de promoção por parte do Governo da RGE para a língua kombe.

3.1.5 Seki

A língua seki (syi), também conhecida como séki, seke, sheke, sekiana, baseke, baseque é uma língua bantu (subgrupo B21) minoritária, sendo falada por cerca de onze mil pessoas, principalmente na região litoral norte da parte continental da RGE e no litoral extremo norte do Gabão. Há pouca literatura linguística específica sobre o povo seki e sobre sua língua (Eberhard; Simons; Fennig, 2022). O primeiro trabalho sobre a língua é o de Raponda Walker (1916/1917), seguida pela lista de palavras de Trilles (1935) e pelas notas gramaticais de González Echegaray (1959) e descrição da fonologia e morfologia de Ondo-Mébiame (1986). Há também algumas menções à língua em trabalhos comparativos dispersos. A língua se mantém robusta na comunidade, embora haja um uso concomitante ao fang (Eberhard; Simons; Fennig, 2022). Não há nenhuma política oficial do Governo da RGE para a língua seki.

3.1.6 Bapuku

A língua bapuku (bnm), também conhecida como papoko, bapuu, batanga, tanga, noho, banôho, banok, naka e puku é uma língua bantu (subgrupo A32) minoritária, sendo falada por cerca de oito mil pessoas na RGE, principalmente ao longo da costa central da parte continental, e na República de Camarões (Eberhard; Simons; Fennig, 2022). Há uma literatura linguística razoável no século XIX sobre o povo bapuku, como por exemplo, Anonymous (1881), e sobre sua língua (Hammarström; Forkel; Haspelmath; Bank, 2020). Há alguns trabalhos linguísticos específicos sobre a língua, como Adams (1907), Siroma (1980), Kouam (1988), Van Hille (1989) e Kouankem (2003), além de menções em trabalhos pan-bantu. Segundo Eberhard; Simons; Fennig (2022), a língua não possui falantes monolíngues, muitos usam o espanhol (com muitos sendo alfabetizados nessa língua), o fang ou o francês em suas trocas comerciais. Adicionalmente, os autores sugerem que a língua se encontra em um processo de obsolescência, com uma diminuição do número de falantes e, por isso, deve estar ameaçada. Não há políticas oficiais de promoção do Governo da RGE para a língua bapuku.

3.1.7 Benga

A língua benga (bng) é uma língua bantu (subgrupo A34) minoritária, sendo falada por cerca de quatro mil pessoas na RGE, principalmente nas ilhas de Corisco, Elobey Grande e Elobey Chico, bem como nos territórios vizinhos no continente, inclusive no Gabão (Eberhard; Simons; Fennig, 2022). Há uma literatura linguística razoável no século XIX e começo do XX sobre o povo benga, como, por exemplo, Mackey (1855), Good (1879), Salvadó y Cos (1891), Mackey; Nassau (1892), Pérez; Sorinas (1928), entre outros, além de uma série de estudos comparativos pan-bantu nos últimos anos. Segundo Eberhard, Simons e Fennig (2022), a língua não possui falantes monolíngues, muitos usam o espanhol (com a maioria dos alfabetizados nessa língua), além de a língua ser usada majoritariamente pelos idosos, principalmente fora da ilha de Corisco. Adicionalmente, os autores sugerem que a língua está ameaçada de extinção, com uma diminuição do número de falantes. Não há políticas oficiais de promoção do Governo da RGE para a língua benga.

3.1.8 Lengue

A língua lengue (bxc), também conhecida como balengue, molengue, molendji é uma língua bantu (subgrupo B21) minoritária, sendo falada por menos de mil e quinhentas pessoas, principalmente na região de Bata na RGE. Há pouca literatura sobre o povo lengue (Eberhard; Simons; Fennig, 2022; Ecuere Dibomo, 1961) e sobre sua língua (De Granda Gutiérrez, 1984b; González Echegaray, 1959). Segundo Eberhard, Simons e Fennig (2022) a língua está ameaçada de extinção, pois o número de falantes está diminuindo e aparentemente não há usuários de lengue como segunda língua, o que indica o seu pouco prestígio. Além disso, como não há falantes monolíngues, as demais línguas poderosas do entorno (fang e espanhol) pressionam os usuários. Não há políticas oficiais de promoção do Governo da RGE para a língua lengue.

3.1.9 Iyasa

A língua iyasa (yko), também conhecida como bongwe, iyassa, maasa, yasa e yassa é uma língua bantu (subgrupo A33), sendo falada por cerca de mil pessoas na região da vila litoral de Edjabe, na RGE (Eberhard; Simons; Fennig, 2022). Ainda segundo o Ethnologue, a língua é também falada na República de Camarões e no Gabão, por aproximadamente duas mil pessoas. Eberhard, Simons e Fennig (2022) afirmam que a língua se encontra ameaçada de extinção, pois somente alguns jovens a utilizam, embora todos os adultos também o façam. Assim como o gyele, as línguas desse grupo, que também é composto por pigmeus da África Equatorial estão ameaçadas pelas línguas poderosas circundantes. Os primeiros trabalhos linguísticos sobre o iyasa foram publicados por Bot (1992; 1997a; 1997b; 1998), seguido por Bouh Ma Sitna (2004) e, recentemente, por um estudo sociolinguístico sobre a manutenção da língua por Belew (2020). Os estudos disponíveis sobre o iyasa, na maior parte das vezes, se concentram na variedade falada nos Camarões. Não há políticas oficiais de promoção do Governo da RGE para a língua iyasa.

3.1.10 Gyele

A língua gyele (gyi), também conhecida como bagyéli, giele, gieli, gyeli, bagyele, bagiele, bajeli, bajele, bogyeli, bogyel, bondjiel, bako, bekoe, bakuele, bakola, bikoya, babinga, baka, likoya é uma língua bantu (subgrupo A801), que é falada por cerca de cinquenta pessoas (somadas a possivelmente algumas centenas em lugares isolados) na RGE, principalmente nas províncias do litoral e na parte centro-sul do país (Eberhard; Simons; Fennig, 2022). Na República de Camarões, contudo, a variante do gyele conhecida como Ngòló possui entre quatro mil e cinco mil falantes e tem sido objeto de um projeto de documentação (Grimm, 2021; Grimm; Ngue Um; Duke, 2020; Grimm *et al.*, 2009-2017). Grimm (2021, p. 2) afirma que o estilo de vida do povo gyele (tradicionalmente caçadores-coletores) com seus padrões de migração de uma vila para a outra, bem como a ausência de dados demográficos oficiais faz com que seja difícil precisar o tamanho da comunidade. Por fim, 'Gyeli speakers are shifting to the languages of their farmer neighbors, a trend which both fragments Gyeli into different dialects and contributes to the language's endangerment.' (Grimm, 2021, p. 2-3), um padrão cada vez mais comum na região. Eberhard; Simons; Fennig (2022) também sugerem que a língua esteja ameaçada de extinção. Para além dos estudos etnográficos, há uma tímida literatura linguística, que se inicia no último quartil do século XX sobre a língua gyele, como por exemplo, Rénaud (1976) e Borchardt (2012) que, como as já citadas, têm como foco a língua falada nos Camarões ou no Gabão. Não há políticas oficiais de promoção do Governo da RGE para a língua gyele.

3.2 Línguas crioulas

Na RGE são faladas duas línguas crioulas: o fa d'Ambô e o pichi. A primeira é falada principalmente na ilha de Ano Bom, e a última majoritariamente em Bioko, mas também nas áreas mais urbanizadas da parte continental do país.

3.2.1 Fa d'Ambô

A língua fa d'Ambô (fab), também conhecida como fa d'Ambu, annobonense, annobonese, e annobonés, é uma língua crioula lexificada pelo português, sendo falada por cerca de seis mil pessoas na RGE, principalmente nas ilhas de Ano Bom e Bioko (Eberhard; Simons; Fennig, 2022), mas com cerca de 600 pessoas na diáspora anobonesa, principalmente na Espanha (Araujo *et al.*, 2013). A língua é geneticamente relacionada às línguas crioulas autóctones de São Tomé e Príncipe (santome (cri), lung'le (pre) e angolar (aoa)). O espanhol é usado amplamente pela população anobonesa, nos domínios oficiais e educacionais e as taxas de letramento em espanhol são altas. Ainda há, no entanto, alguns falantes monolíngues de fa d'Ambô (Agostinho, 2021). A língua foi inicialmente descrita no século XIX por Vila (1891) e Barrena (1957)¹⁰. Posteriormente, no século XXI, surgiram os trabalhos de Zamora Segorbe (2010), Post (2013), Bandeira (2017), Agostinho; Araujo; Santos (2019), Hagemeijer; Maurer-Cecchini; Zamora Segorbe (2020), Agostinho (2021), além de menções em trabalhos que discutem as características linguísticas dos crioulos lexificados pelo português do Golfo da Guiné. A língua é vigorosa (EGIDS 6a)¹¹ e todos os anoboneses a utilizam em Ano Bom (Agostinho (2021). A partir do nosso trabalho de campo realizado em 2012, observamos que o fa d'Ambô é considerado um elemento de resistência política e cultural em Ano Bom. Não há políticas oficiais de promoção do Governo da RGE para a língua fa d'Ambô.

3.2.2 Pichi

O pichi (fpe), também conhecido como pichinglis, Fernando Po Creole English, Equatorial Guinean Pidgin, criollo, fernandino, Fernando Po krio e pidgin da Guiné Equatorial, é uma língua crioula lexificada pelo inglês – fortemente baseada no krio (kri), que tem sido levado à Guiné Equatorial por trabalhadores migrantes de Sierra Leoa desde o segundo quartel do século XIX (Fyfe, 1962) –, sendo falada por mais de cem mil pessoas na RGE, principalmente na ilha de Bioko, tanto como primeira, como segunda língua. Dessa forma, ao mesmo tempo em que o pichi é a segunda opção de falantes de outras línguas da RGE, essas outras línguas também são utilizadas pelos falantes de pichi como primeira língua, além do inglês e do espanhol (Eberhard; Simons; Fennig, 2022). Embora o seu nome seja associado a um pidgin, trata-se de uma língua crioula, uma vez que há falantes de pichi como primeira língua (além de falantes monolíngues) e que é a língua de comunidades específicas e associadas linguística e culturalmente. O pichi é usado amplamente pela população na Ilha de Bioko e, cada vez mais, está presente nas comunidades urbanas do continente, porém afastada dos domínios oficiais e educacionais. No século XX, foram publicados dois dicionários do pichi (De Zarco, 1938; Mangado, 1919) e Yakpo (2009; 2010; 2013; 2019) tem publicado gramáticas da língua em inglês e em espanhol. Como língua urbana, o pichi tem se tornado a principal língua veicular na RGE. A língua é vigorosa e, assim como o fang, tem se tornado uma ameaça às línguas étnicas minoritárias. Não há políticas oficiais de promoção do Governo da RGE para a língua pichi.

¹⁰ O reverendo claretiano Natalio Barrena (1867-1925) elaborou a sua gramática possivelmente no fim do século XIX e começo do XX, embora seu trabalho tenha sido publicado postumamente em 1957.

¹¹ A língua é usada oralmente por todas as gerações, é aprendida pelas crianças e as condições para a oralidade sustentável são atendidas. Para mais informações, ver <https://www.ethnologue.com/methodology/#languageStatus>.

3.3 Línguas europeias

Na RGE, três línguas europeias possuem o estatuto de idioma oficial: o espanhol (oficial desde a constituição de 1982), o francês (desde 1989) e o português (desde 2010). Assim, a RGE é o único país africano que possui três línguas europeias como oficiais. O espanhol (esp) tem sido usado na administração pública desde a metade do século XIX, quando o Reino de España efetivamente passou a controlar o território. O francês (fra) foi oficializado em 1989 com o objetivo de inserir a RGE na área econômica francófona da África Atlântica. Por fim, o português (por) recebeu o status de língua oficial em 2010, como requisito parcial para a admissão da RGE na Comunidade de Países de Língua Portuguesa (CPLP).

3.3.1 Espanhol

A língua espanhola é uma língua românica falada por mais de um milhão de pessoas na RGE. Contudo, não há consenso sobre o número de falantes de espanhol como primeira língua. Eberhard, Simons e Fennig (2022), citando o Instituto Cervantes, estimam haver 905 mil nativos. No entanto, esse número parece exagerado. De acordo com o quadro 1 do documento *El español: una lengua viva. Informe 2022* (Instituto Cervantes, 2022) há na RGE uma população de 1.454.789 indivíduos (citando dados do *Instituto Nacional de Estadística* da RGE de 2022), dos quais 74%, ou seja, 1.076.544, formariam o 'grupo de dominio nativo', ou seja, 'En el CDN se contabilizan los bilingües como hispanohablantes, pero no los monolingües en otras lenguas' (Instituto Cervantes, 2022, Cuadro 1, Nota 3), ao passo que 378.245 pessoas (26%) formariam o 'grupo de competencia limitada', isto é,

El GCL incluye a los hablantes de español de segunda y tercera generación en comunidades bilingües, a los usuarios de variedades de mezcla bilingües y a las personas extranjeras de lengua materna diferente del español residentes en un país hispanohablante (Instituto Cervantes, 2022, Cuadro 1, Nota 4).

Portanto, segundo o documento, todos os habitantes da RGE seriam falantes de espanhol, nativos ou falantes de segunda língua. Em um país com baixas taxas de eficiência escolar, grande parte do território coberta por florestas, vários grupos étnicos semi-nômades, capacidade reduzida do Estado de oferecer os serviços sociais mínimos, a afirmação do documento do Instituto Cervantes pode estar exagerada. À título de comparação, Nistal Rosique (2007, p. 74), por sua vez, afirma que 74% da população aprende a língua na escola e, portanto, tais pessoas devem ser consideradas falantes de espanhol como segunda língua:

En su artículo incluido en el anuario del Instituto Cervantes *El español en el mundo* 2005, Trinidad Morgades, vicerrectora de la Universidad Nacional de Guinea Ecuatorial, indica que hay tres clases de hablantes de español en el país: los que lo hablan y lo escriben adecuadamente, los que lo han aprendido como segunda lengua y los que necesitan programas de alfabetización. Nosotros añadiríamos que en el primer grupo se incluirían fundamentalmente los mayores de cuarenta años, que porcentualmente supondrían entre un 10 % y un 15 % de la población, mientras que en el segundo grupo estarían la gran mayoría de los hablantes guineoecuatorianos, con un porcentaje que rondaría el 74 %, dejando para el último grupo, el de los hablantes que necesitan programas de alfabetización, entre un 12 % y un 13 % de la población. Nistal Rosique (2007, p. 74, grifos nossos).

Destarte, o espanhol goza de um estatuto importante, é lecionado nas escolas de todos os níveis e está presente na documentação oficial e privada. Apesar de a maior parte da população falar espanhol, 'there is no empirical study or language survey/data which has determined the level of proficiency or the number of monolingual speakers among Equatorial Guineans' (Gomashie, 2019, p. 3). Há vários estudos sobre a variedade espanhola na RGE. Desde 1951, os estudiosos têm destacado as características da variedade equato-guineense e atitudes locais em relação à língua, entre eles, González Echegaray (1951), Castillo Barril (1964; 1969), De Granda Gutiérrez (1984a; 1984b; 1985; 1994, entre outros), Lipski (1985; 2002a; 2002b; 2004; 2007; 2008); (Lipski, 2014), Bibang Oyee (2002; 2014), Chirila (2015), Schlumpf (2016), (Gomashie, 2019), entre tantos outros. Por ser a primeira língua oficial do país, os atos de Estado de Governo, bem como a escolarização e a maior parte da mídia são em espanhol.

3.3.2 Francês

A língua francesa é uma língua romântica falada por cerca de trezentos e oitenta mil pessoas na RGE, oficial desde 1998. Gomashie (2019) apresenta as razões para a oficialização do francês:

As a result of Equatorial Guinea seeking closer economic ties with Francophone countries in 1983 and its membership in the Economic Community of Central African States in 1985, French was incorporated, in 1988, in the educational system as a compulsory subject and an official national language in 1997 (Gomashie, 2019, p. 2).

Assim, como uma importante língua no comércio regional com os países de língua oficial francesa da região, e no comércio e atividades diárias locais com os imigrantes e trabalhadores migrantes de países da região, bem como segunda língua em várias comunidades linguísticas nacionais cujas línguas se espalham pelo Camarões e Gabão, a língua francesa representa um importante papel de língua franca no território. Igualmente, é ensinada nas escolas primárias e secundárias do país (Eberhard; Simons; Fennig, 2022).

3.3.3 Português

A língua portuguesa é uma língua romântica, oficial na RGE desde 2010. A oficialização da língua está ligada ao desejo do Governo Obiang em fazer parte da CPLP. Em 2006, a RGE foi admitida, como membro-observador da CPLP, uma organização internacional multilateral cujo objetivo é promover a cooperação entre os países-membros. Posteriormente, em 2010, o país solicitou à CPLP o status de membro pleno. Para isso, segundo a CPLP, era necessário que a língua portuguesa fosse declarada oficial e que o país provasse uma sólida ligação histórica com a língua portuguesa (Araujo; Agostinho; Christofolletti; Freitas *et al.*, 2013, p. 30, nota 12). De fato, o fado e a religião católica na ilha de Ano Bom são provas incontestes da ligação com o mundo português. Em 2014, a RGE foi aceita como membro pleno. Porém, o português não é efetivamente falado no país, exceto por uma pequena comunidade de imigrantes e expatriados, missionários, funcionários públicos e privados, ligados às embaixadas e às empresas que atuam na região. Embora seja uma língua oficial, há poucas e difusas tentativas de promover o português por parte do Governo da RGE. Igualmente, o português sequer empregado na comunicação oficial. Portanto, trata-se de uma de oficialidade praticamente figurativa.

4 Discussão

A complexidade de se definir os termos ‘política linguística’ e ‘planejamento linguístico’ decorre da natureza das abordagens científicas nos últimos sessenta anos, durante a fundação e desenvolvimento dessas áreas de pesquisa (Tollefson; Pérez-Milans, 2018). De um lado, há as atividades de planejamento e política linguística *top-down*, isto é, os discursos e as práticas dos entes organizados do Estado (ou por ele delegados) para interferir no capital linguístico de uma ou mais de uma língua, fazendo com que essa intervenção tenha impacto (para incorporar ou repelir uma determinada língua) no conceito de nação e nacionalismo, integração política e social dos cidadãos, políticas educacionais de alcance geral, promoção de documentação, produção de materiais didáticos e paradidáticos, escolarização, desenvolvimento econômico e conexão a parceiros, regionais e globais utilizando uma ou mais de uma língua (Bourdieu, 1991). Por outro lado, devemos também considerar as atividades de planejamento e política linguística *bottom-up*, ou seja, aquelas que partem da atuação de indivíduos, famílias, instituições não-governamentais e universidades (sem mandato do Estado para essa função) para promover a documentação e o uso de uma língua em um ambiente, seja essa promoção organizada ou não (Johnson, 2018). Assim, na RGE, podemos observar uma intervenção governamental ao estabelecer suas línguas oficiais (o espanhol, o francês e o português), ora olhando para o passado (o espanhol e seu legado como potência imperialista), ora para o presente (o francês e sua realidade inconteste no tecido político regional e de migração), ora para o futuro (o português como via de acesso a oportunidades econômicas e mercados de um grupo de países), muito embora, na prática, o governo tenha mostrado pouca intenção de intervir, na documentação, na difusão e na promoção das línguas francesa e portuguesa no sistema educacional e cultural como um todo¹². Essa pouca capacidade ou vontade pode ser observada, por exemplo, nos investimentos em educação. Segundo os dados do Banco Mundial, em 2020, a RGE aplicou 1% do seu produto interno bruto na área, ao passo que seus vizinhos, Gabão (2,7%) e Camarões (3,1%) investiram mais (Worldbank, 2020). Igualmente, podemos observar a intervenção e atuação dos falantes ao escolherem usar uma língua, mesmo que essa sofra preconceito e seja alvo da repulsa dos entes do Estado e de membros da sociedade, como é o caso do pichi na RGE (Yakpo, 2011) ou de qualquer uma outra língua não-europeia, muito embora, em geral, as línguas minoritárias possuam prestígio em suas comunidades. Adicionalmente, atividades de legitimação, através de documentação e difusão na mídia (rádio e internet) são cruciais para as línguas marginalizadas. Por fim, há a atuação dos falantes ao privilegiarem uma determinada língua, seja na forma de uma substituição de línguas (*language shift*) ou a resistência linguística ao continuarem a usar línguas marginalizadas.

¹² Portanto, aqui não nos limitamos à definição de Crystal (2008) segundo a qual o planejamento linguístico inclui a documentação das línguas e o desenvolvimento de atos que promovam sua escolha e uso, a preparação de instrumentos como grafia, desenvolvimento de literaturas, apoio a programas educacionais e treinamento de recursos humanos em diferentes níveis para as escolas. Para Johnson (2013), as políticas linguísticas causam impacto na oficialização das línguas, com influências nos sistemas econômicos, políticos e educacionais. Igualmente, a circulação e o capital linguístico de uma língua atuam em mecanismos de poder e empoderamento, levando à atuação, seja de agentes governamentais, seja do indivíduo ou da família/comunidade de fala.

Posto que a RGE possui línguas com menos de cinquenta mil falantes – bube (50 mil), kombe (16 mil), kwasio (15 mil), seki (11 mil), bapuku (8 mil), fa d'Ambô (6 mil), benga (4 mil), lengue (1,5 mil), iyasa (1 mil) e gyle (menos de 100) – a pressão que essas línguas minoritárias e seus falantes sofrem das línguas francas (fang e pichi), do espanhol e do francês é enorme. Por isso, não é incomum que as pessoas abandonem as suas línguas étnicas, ou seja, deixem de transmiti-las às crianças e optem por outras. Isso acontece quando a viabilidade econômica da região está associada a uma língua, ou no contexto de migração, e/ou quando uma língua mais poderosa se apresenta como viável no ensino, na mídia e nas relações interpessoais. Populações pequenas costumam ser as mais vulneráveis, principalmente quando estiverem em contextos urbanos. Ao mesmo tempo, quando a comunidade perde a confiança na língua, por achar que as gerações mais jovens, em situações de multilinguismo, incorporam muitos elementos, ao ponto de impedir ou atrapalhar a comunicação com os falantes mais idosos, ou até mesmo quando os jovens alegam não ser mais proficientes. Muitas vezes, por ameaça de uso de violência, os adultos forçam os jovens a substituírem a língua étnica minoritária por uma língua majoritária. Por fim, discursos preconceituosos podem ter um efeito negativo na língua, porém, como elemento de resistência, a comunidade pode ser refratária a esses discursos. Na RGE, o pichi tem ampliado o seu número de falantes na capital, apesar de discursos agressivos contra a língua:

Apesar de su gran importancia como lengua de comunidad y lengua franca (inter)nacional, el pichi carece de reconocimiento y apoyo oficial, está conspícidamente ausente del discurso público y de los medios oficiales, y no tiene presencia en la política educativa del país. Por consiguiente, debido a esta particular historia, el pichi epitomiza el rechazo político a las lenguas africanas, no sólo en Guinea Ecuatorial, sino también en África en general (Yakpo, 2011, p. 17-18).

A oficialização de uma língua na RGE não implica, necessariamente, em sua promoção. A oficialização do espanhol na Constituição de 1982 é a exceção, juntamente com o francês, por sua vez, devido ao contexto regional, pois os países circunvizinhos da RGE, Camarões e Gabão, têm a língua como oficial. Ademais, há dezenas de milhares de trabalhadores migrantes e suas famílias falantes de francês no país. Após o término do governo Macías, milhares de equato-guineenses (auto)exilados em países de língua oficial francesa na região retornaram à RGE. Portanto, adotar o francês, além de permitir ao país uma inserção em uma área econômica importante, permite fazer justiça aos usos linguísticos de grande parte da população residente. Finalmente, a oficialização do português pode ser considerada como um capricho diplomático do Governo. Ainda assim, foi a partir da oficialização do português que a pressão da CPLP se fez importante para que a pena capital, como forma de punição do sistema jurídico do país, fosse paulatinamente abandonada.

Incluindo o espanhol e o francês, ambos presentes no sistema escolar da RGE e o primeiro sendo a língua dos atos do Estado e do Governo, a história do planejamento e das políticas linguísticas do país pode ser dividida em quatro fases:

- a. escolarização colonial missionária e, posteriormente, laica, em espanhol
- b. documentação por missionários (até 1950) e linguistas até 1968
- c. documentação por linguistas e organizações não-governamentais pós-1968
- d. escolarização em massa em espanhol; francês (como língua estrangeira)

A fase da escolarização colonial missionária e, posteriormente, laica, em espanhol, se inicia quando o Reino da Espanha começa a colonização, a partir da ilha de Bioko, após a saída dos ingleses. A colonização e a 'aculturação' do território continental só terá início na década de 1930 (Negrín Fajardo, 1993, p. 16). O objetivo da ditadura franquista, ainda segundo Negrín Fajardo (1993), era delegar a tarefa da educação colonial a missionários:

[n]el caso de Guinea, que durante una larga etapa fue prácticamente territorio de control misionero hasta que las autoridades civiles se fueron haciendo paulatinamente con el dominio real de la zona. No obstante, hasta la independencia de Guinea, la influencia de las misiones fue siempre muy importante (Negrín Fajardo, 1993, p. 16).

Na GE colonial, 'la administración con los misioneros a la cabeza, combate con rigor la cultura y tradición bantú y la educación espontánea ancestral.' (Negrín Fajardo, 1993, p. 46). Naturalmente, a maior barreira às políticas educacionais, sejam missionárias, sejam oficiais, era o fato de a população simplesmente não falar espanhol e nem conhecer os rudimentos da 'cultura espanhola'. De qualquer forma, havia conflitos entre os administradores do Governo (que desejavam viabilizar econômica e socialmente a colônia) e os missionários que desejavam 'convertir a los paganos y erradicar la poligamia y la promiscuidad sexual.' Negrín Fajardo (1993, p. 52). A situação de conflito teve fim com a II República espanhola que implementou um ensino laico. O regime de Franco promoveu na GE uma política de aculturação, com avanços na educação básica, bem como na formação de quadros dirigentes locais, segundo os padrões imperiais. Para os locais, essa política representou desvalorização das culturas e línguas do país. Ao mesmo tempo, o regime franquista implanta a transmissão via rádio no país em 1952-3, que serve com um meio para difundir a 'cultura y civilización hispánica' na GE (Negrín Fajardo, 1993, p. 148). Ao lado da educação básica, houve também um forte investimento em escolas de ensino médio e profissionalizantes públicas e algum investimento privado (principalmente religioso). O ensino superior era realizado na metrópole (Pélissier, 1964).

Até 1968, o *Instituto de Estudios Africanos*, do *Consejo Superior de Investigaciones Científicas*, em Madrid, promoveu a investigação etnográfica e linguística com a publicação de livros e do periódico *Archivo del Instituto de Estudios Africanos*, descontinuado em 1966 (Negrín Fajardo, 1993, p. 13). Com a independência, o trabalho de documentação linguística passa a ser isolado, concentrando-se em dissertações de mestrado, teses de doutorado ou trabalhos de compilação pan-bantu, realizados por pesquisadores em universidades, quase sempre da Europa, dos Estados Unidos ou do Senegal.

Após a independência, há um hiato nas publicações sobre as línguas da RGE. Isso se deve ao caráter isolacionista do governo Macías e, inicialmente, às dificuldades impostas a acadêmicos pelo regime de Obiang. Ao mesmo tempo, a tarefa de fazer trabalho de campo na GE não é simples, com alto custo e riscos inerentes ao trabalho em regiões de extrema pobreza, como acesso à água e à energia, emprego, questões de segurança alimentar, saúde etc. Nos últimos vinte anos, os trabalhos sobre as línguas da RGE se limitam a descrições e análises do espanhol, do pichi, do fang e do fa d'Ambô.

Por fim, apesar de alguns problemas, o governo promove a escolarização em massa em espanhol e do francês, como segunda língua. O *Centro Cultural de España* do Instituto Cervantes, além de promover o espanhol, promove cursos da língua fang. Em 2022, foi inaugurada a Tribuna do Hispanismo Ecuatoguineano, que pretende agregar especialistas no espanhol da RGE¹³. Porém, a documentação e a investigação sobre as demais línguas são realizadas por acadêmicos, de forma isolada, ou simplesmente não são realizadas.

5 Considerações finais

O objetivo desse texto foi descrever e analisar a situação linguística na RGE, fazendo um levantamento das políticas oficiais passadas e presentes, desde a tomada do território pela Espanha, até os dias atuais. A língua do Estado, da mídia e da escolarização é o espanhol. Aprendido na escola, é uma língua dominada pela maior parte da população como língua estrangeira. Assim, o espanhol é a língua nacional no sistema oficial, e o francês é a língua do comércio regional e dos expatriados. Mostramos também que, embora haja três línguas oficiais, o português tem uma existência *in iure*, ou seja, apenas protocolar, pois não há uma comunidade relevante falante de português e as relações econômicas com os países de língua portuguesa são mínimas, mesmo que haja (fisicamente) um país de língua portuguesa, São Tomé e Príncipe, entre as ilhas de Bioko e Ano Bom. Todavia, a pouca relevância de São Tomé e Príncipe no comércio internacional reflete também na sua invisibilidade para a RGE. O francês, por seu turno, é a língua dos países vizinhos Gabão e Camarões, bem como é utilizada pelo crescente número de trabalhadores importados dessas regiões. Assim, o francês se constitui como um importante elemento na economia da região.

Para além dessas línguas oficiais, a sobrevivência ou a obsolescência das demais línguas em suas comunidades de fala está ligada a uma série de fatores. Embora todas as línguas étnicas possuam prestígio em suas comunidades, a manutenção de uma língua depende de outros fatores. Assim, uma língua regional forte, seja no número de falantes, no seu *soft-power*, ou no seu capital linguístico, atua como um fator de desestabilização de uma língua minoritária. Igualmente destacamos, de um lado, o crescimento da língua pichi, principalmente nas aglomerações urbanas como a língua da juventude e dos trabalhadores estrangeiros oriundos de regiões de fala inglesa e, por outro lado, do fang, devido ao tamanho de sua população falante e igualmente por ser a língua do grupo politicamente dominante nos últimos cinquenta anos. Para além da atuação das línguas oficiais, o fang e o pichi têm atraído cada vez mais falantes, por diversas razões. O fang por ser uma língua falada por um grupo étnico numeroso e poderoso politicamente e, o pichi, por ser uma língua urbana de Bioko falada por uma população jovem. Ao mesmo tempo, a incapacidade ou o desinteresse do governo em promover a documentação e a difusão das línguas étnicas contribui para o quadro de ameaça às línguas minoritárias. Dessa forma, não apresentar uma política linguística de promoção das línguas crioulas e étnicas é, em si, uma tomada de posição política. O apagamento dessas línguas no discurso oficial é feito com um propósito político de silenciar a variedade linguística e cultural dos povos. Ao mesmo tempo, a maior ameaça a todas as línguas étnicas é a falta de reconhecimento e valorização por parte do governo.

¹³ Disponível em https://www.cervantes.es/sobre_instituto_cervantes/prensa/2022/noticias/observatorio-hispanismo-guinea-ecuatorial.htm, acessado em 05/11/2022.

guas faladas na RGE provém da situação econômica da maior parte da população, que vive na extrema pobreza e sem perspectivas. É difícil promover a preservação das línguas maternas quando grande parte da população não tem acesso à água tratada, à energia elétrica e a um sistema de alimentos, a um sistema de saúde e de educação dignos.

Nos últimos anos, o *Centro Cultural de España* até atuado na promoção das línguas minoritárias, principalmente do fang. Ao lado dele, alguns linguistas têm trabalhado na documentação e divulgação dessas línguas, porém com uma sociedade civil não organizada, há pouco ou nenhum instrumento de pressão no país para que o quadro aqui descrito seja modificado e que as línguas étnicas sejam promovidas na RGE. Ademais, os poucos trabalhos sobre essas línguas tendem a ser direcionados à comunidade científica, sem objetivos pedagógicos, com exceção do fang. Por fim, o diagnóstico preciso de Yakpo (2011) para a promoção e documentação das línguas locais não encontrou, infelizmente, acolhida no sistema educacional e na mídia da RGE.

Referências

- ABAD, I. *Elementos de la gramática bubi*. Madrid: Editorial del Corazón de María, 1928.
- ADAMS, G. A. Die Sprache der Banôho. *Mittheilungen des Seminars für Orientalische Sprachen*, Berlin, v. 10, p. 34-83, 1907.
- AGOSTINHO, A. L. Fa d'Ambô (Equatorial Guinea). *Language Documentation and Description*, Charlottesville, v.20, p. 123-134, 2021. DOI: <https://doi.org/10.25894/ldd39>.
- AGOSTINHO, A. L.; ARAUJO, G. A. D.; SANTOS, E. F. D. Interrogative particle and phrasal pitch-accent in polar questions in Fa d'Ambô. *Boletim do Museu Paraense Emílio Goeldi Ciências Humanas*, Belém, v.14, n. 3, p. 1-16, 2019. DOI: <https://doi.org/10.1590/1981.81222019000300009>.
- AGOSTINHO, A. L.; BANDEIRA, M.; ARAUJO, G. A. D. O lung'le na educação escolar de São Tomé e Príncipe. *Trabalhos em Linguística Aplicada*, Campinas, v.55, n. 3, p. 591-618, 2016. DOI: <https://doi.org/10.1590/010318135164183401>.
- ÁLVAREZ-CHILLIDA, G.; NERÍN, G. Introducción. Guinea Ecuatorial: el legado de la colonización española. *Ayer*, Valencia, v. 109, n. 1, p. 13-32, 2018. DOI: <https://doi.org/10.55509/ayer/109-2018-01>.
- ANONYMOUS. Notes on the People of Batanga. *The Journal of the Anthropological Institute of Great Britain and Ireland*, London, v. 10, p. 463-470, 1881. DOI: <https://doi.org/10.2307/2841552>.
- ARAUJO, G. A. D.; AGOSTINHO, A. L.; CHRISTOFOLETTI, A.; FREITAS, S. et al. Fa d'ambô: língua crioula de Ano Bom. *Cadernos de Estudos Linguísticos*, Campinas, v. 55, n. 2, p. 25-44, 2013. DOI: <https://doi.org/10.20396/cel.v55i2.8637289>.
- AYMEMI, A. *Diccionario español-bubi*. Madrid: Editorial del Corazón de María, 1928.
- BANDEIRA, M. *Reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné*. München: Lincom, 2017.
- BARRENA, N. *Gramática anobonesa*. Madrid: Instituto de Estudios Africanos, 1957.
- BATES, G. L. *A Grammar of the FANG language as spoken on the Como and Benito River*. s.l.: [s.n.], 1899. p. 115.

- BEBAN, C. S.; ATINDOGBE, G. G.; DOMCHE, E.; BOT, D. M. L. Classification of the languages of Cameroon and Equatorial Guinea on the basis of lexicostatistics and mutual intelligibility. *African Study Monographs*, I., v. 28, n. 4, p. 181-204, 2007.
- BELEW, A. K. *Sociolinguistic documentation of language shift and maintenance in lyasa*. 2020. 185 f. Tese (Doutorado em Linguística) — Departamento de Linguística, University of Hawai'i, Mānoa. 2020.
- BIBANG OYEE, J. *Curso de Lengua Fang*. Malabo: Centro Cultural Hispano-Guineano Ediciones, 1990.
- BIBANG OYEE, J. *El Español Guineano—Interferencias, Guineanismos*. Malabo: Julián Bibang Oyee, 2002.
- BIBANG OYEE, J. *Diccionario español-fang / fang-español*. Madrid: ACAL/AECID, 2014.
- BOLEKA, J. B. *Aspectos lingüísticos y sociolingüísticos del bubi del noroeste en relación con el castellano y el francés de los países francófonos del área ecuatorial*. Dissertação (Mestrado em Linguística) — Escuela de Doctorado, Universidade Complutense de Madrid, 1986.
- BOLEKA, J. B. El aumento o actualizador definido en lengua bubi. *Muntu: revue scientifique et culturelle du CICIBA (Centre International des Civilisations Bantu)*, Libreville, v. 7, p. 179-199, 1987.
- BOLEKA, J. B. *Curso de lengua bubi*. Malabo: Centro Cultural Hispano-Guineano, 1991.
- BOLEKA, J. B. *Diccionario español-bubi = Ŋribúkku ra balláa békobé-léepanná*. Madrid: Akal, 2009.
- BORCHARDT, N. The verbal system in Gyeli. In: Humboldt-Universität zu Berlin Linguistisches Kolloquium, 2012, Berlin. November 20, 2012. p. 20. Disponível em: <https://www.iaaw.hu-berlin.de/de/afrika/linguistik-und-sprachen/veranstaltungen/afrikalinguistischeskolloquium/papers-wintersemester-2012-13/nadine-borchardt-the-verbal-system-in-gyeli>
- BOT, D. M. L. *Phonologie générative du Yasa*. 1992. 274 f. Dissertação (Mestrado em Língas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé I, Yaoundé, 1992.
- BOT, D. M. L. La nominalisation en yasa. *Afrikanistische Arbeitspapiere*, [s.l.], v. 52, p. 19-28, 1997a.
- BOT, D. M. L. Structure syllabique et lois morphémiques du yasa. *Afrikanistische Arbeitspapiere*, Köln, v. 49, p. 31-43, 1997b.
- BOT, D. M. L. Temps verbaux et aspects du Yasa. *Afrikanistische Arbeitspapiere*, Libreville, v. 53, p. 47-65, 1998.
- BOUH MA SITNA, C. L. *Le Syntagme Nominal du Yasa*. 2004. 94 f. Dissertação (Mestrado em Língas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé I, Yaoundé, 2004.
- BOURDIEU, P. *Language and symbolic power*. Cambridge: Harvard University Press, 1991.
- BURTON, F. R. *Wanderings in West Africa*. Mineola: Dover Publications Inc., 1992.
- CALDEIRA, A. M. Uma ilha quase desconhecida. Notas para a história de Ano Bom. *Studia Africana*, Porto, v. 17, p. 99-109, 2006.
- CASTILLO BARRIL, M. El español en la Guinea Ecuatorial. *Español Actual*, Madrid, v. 3, p. 8-9, 1964.
- CASTILLO BARRIL, M. La influencia de las lenguas nativas en el español de Guinea. *Archivo de Estudios Africanos*, Madrid, v. 20, p. 46-71, 1969.
- CASTILLO-RODRÍGUEZ, S. The first missionary linguistics in Fernando Po: Transliteration and the quest of Spanishness in an Anglicized colony. In: KLAUS, Z. e BIRTE, K.-R. (eds.). *Colonialism*

and Missionary Linguistics. Berlin/ München/ Boston: De Gruyter, 2015. p. 75-106. DOI: <https://doi.org/10.1515/9783110403169.75>.

INSTITUTO CERVANTES. El español: una lengua viva. Informe 2022. Madrid, 2022. Disponível em: <https://cvc.cervantes.es/lengua/anuario/anuario_22/informes_ic/p01.htm>

CHIRILA, E. M. *Identidad Lingüística en Guinea Ecuatorial: Diglosia y Actitudes Lingüísticas ante el Español*. 2015. 122 f. Dissertação (Mestrado em Línguas Estrangeiras) — Faculdade de Humanidades, Universitat Bergensis, Bergen.

CIA. The World Factbook. Washington, D. C.: [s.n.], 2021.

CLARKE, J. *The Adeeyah vocabulary for the use of schools in Western Africa*. Falmouth, Jamaica: Baptist Mission, 1841.

CLARKE, J. *Sentences in the Fernandian tongue*. Bimbia: Dunfermline Press, 1846.

CLARKE, J. *Introduction to the Fernandian tongue*. Berwick-on-Tweed, D. Cameron, 1848.

CRAVEN, M. Between law and history: the Berlin Conference of 1884-1885 and the logic of free trade. *London Review of International Law*, London, v. 3, n. 1, p. 31-59, 2015. DOI: <https://doi.org/10.1093/lrl/lrv002>.

CRESPI, L. En busca de un enclave esclavista. La expedición colonizadora a las islas de Fernando Poo y Annobon, en el Golfo de Guinea. (1778 – 1782). *Estudios Históricos*, Montevideo, v. 4, p. 1-34, 2010.

CRYSTAL, D. *Language Death*. Cambridge: Cambridge University Press, 2002.

CRYSTAL, D. *Dictionary of linguistics and phonetics*. 6 ed. Oxford: Blackwell, 2008.

DAS NEVES, C. A. A reacção dos habitantes de Fernando Pó e Ano Bom à dominação estrangeira. *Studia*, [s.l.], v. 50, p. 199-214, 1991.

DE GRANDA GUTIÉRREZ, G. Fenómenos de interferencia fonética del fang sobre el español de Guinea Ecuatorial: consonantismo. *Anuario de Lingüística Hispánica*, Madrid, v. 1, p. 95-114, 1984a.

DE GRANDA GUTIÉRREZ, G. Perfil lingüístico de Guinea Ecuatorial. In: *Homenaje a Luis Flórez: estudios de historia cultural, dialectología, geografía lingüística, sociolingüística, fonética, gramática y lexicografía*. Bogotá: Instituto Caro y Cuervo, 1984b.

DE GRANDA GUTIÉRREZ, G. *Estudios de lingüística afro-románica*. Valladolid: Universidad de Valladolid, 1985.

DE GRANDA GUTIÉRREZ, G. (ed.). *Español de América, español de África y hablas criollas hispánicas*. Madrid: Credos, 1994.

DE WULF, M. Population Pyramid. 2024. Disponível em: <<https://www.populationpyramid.net>>.

DE ZARCO, M. *Dialecto inglés-africano o broken-english de la colonia española del Golfo de Guinea*. Turnhout: H. Proost, 1938.

DIEU, M. Les consonnes du ngumba: recherche en phonologie générative. *Bulletin de l'ALCAM (Atlas linguistique du Cameroun)*, Yaoundé, v. 1, p. 33-205, 1976.

EBERHARD, D. M.; SIMONS, G. F.; FENNIG, C. D. (ed.). *Ethnologue: Languages of the World*. Twenty-fifth edition. Dallas: SIL International, 2022.

- ECUATORIAL, Guiné. Retos y mecanismos de protección de la cultura ndowé. In: MEMBA, J. (ed.). *Contando Guinea Ecuatorial*. Bata: ASAMA, 2018.
- ECUERE DIBOMO, B. Notas sobre los balengues. *La Guinea Española*, Malabo, v. 57, p. 273-275, 1961.
- ELLA , E. M. *A Theoretical Model for a Fang-French-English Specialized Multi-Volume School Dictionary*. 2007. 333 f. Tese (Doutorado em Literatura (Lexicografia) — Faculdade de Artes e Ciências Sociais, Stellenbosch University, Stellenbosch, 2007.
- ELIMELECH, B. Noun tonology in Kombe. In: HYMAN, L. M. (ed.). *Studies in Bantu tonology*. Los Angeles: Department of Linguistics, University of Southern California, 1976. p. 113-130.
- ESPAÑA, G. D. Tratado del Pardo de 24 de marzo de 1778 entre España y Portugal de amistad, garantía y comercio. Madrid: Secretaría de Estado y del Despacho de Estado (España), s.d.
- FERNANDEZ, G. L. *Diccionario Español-Kômbè*. Madrid: Instituto de Estudios Africanos, 1951.
- FYFE, C. *A history of Sierra Leone*. Oxford: Oxford University Press, 1962.
- GARCIA CANTÚS, M. A. D. *Fernando Poo: una aventura colonial española en el África occidental, 1778-1900*. 2004. 702p. Tese de Doutorado em História) — Departament d'Història Contemporània, Universitat de València. 2004.
- GELABERT, P.; FERRANDO-BERNAL, M.; DE-DIOS, T.; MATTORRE, B. et al. Genome-wide data from the Bubi of Bioko Island clarifies the Atlantic fringe of the Bantu dispersal. *BMC Genomics*, [s.l.], v. 20, n. 1, p. 179, 2019. DOI: <https://doi.org/10.1186/s12864-019-5529-0>.
- GOBIERNO-DE-GUINEA-ECUATORIAL. Proyecto de Ley Constitucional. 2010.
- GOBIERNO-DE-GUINEA-ECUATORIAL. Ley Fundamental de Guinea Ecuatorial. Malabo: Imprenta del Gobierno Guinea Ecuatorial,
- GOMASHIE, G. A. Language Vitality of Spanish in Equatorial Guinea: Language Use and Attitudes. *Humanities*, [s.l.], v. 8, n. 1, p. 33, 2019. DOI: <https://doi.org/10.3390/h8010033>.
- GONZÁLEZ ECHEGARAY, C. Notas sobre el Español en África Ecuatorial. *Revista de Filología Española*, Madrid, v. 35, p. 106-118, 1951.
- GONZÁLEZ ECHEGARAY, C. *Estudios Guineos*. Madrid: Consejo Superior de Investigaciones Científicas, 1959.
- GONZÁLEZ ECHEGARAY, C. *Morfología y Sintaxis de la lengua Bujeba*. Madrid, Instituto de Estudios Africanos, Consejo superior de investigaciones científicas, 1960. 226 p.
- GOOD, A. I. *Dictionary of the English and Benga languages*. New York: Mission House, 1879.
- GRIMM, N. *A grammar of Gyeli*. Berlin: Language Science Press, 2021.
- GRIMM, N.; NGUE UM, E.; DUKE, D. A documentation of the Bagyeli/Bakola forest foragers of Cameroon. Nijmegen: MPI, The Language Archive 2020.
- GRIMM, N.; NGUE UM, E.; LORENZ, C.; DUKE, D. et al. Bagyeli/Bakola. : The Language Archive 2009-2017.
- HAGEMEIJER, T.; MAURER-CECCHINI, P.; ZAMORA SEGORBE, A. *A grammar of Fa d'Ambô*. Berlin: Mouton de Gruyter, 2020.
- HAMMARSTRÖM, H.; FORKEL, R.; HASPELMATH, M.; BANK, S. Glottolog database 4.2.1 (v4.2.1). 2020.

- IKUGA EBOMBEBOMBE, A. *Cómo se habla y se escribe el kombe*. Barcelona: Institut Catalunya África, 1973a.
- IKUGA EBOMBEBOMBE, A. *Cómo se habla, se escribe y se lee el ndowe*. Barcelona, Institut Catalunya África, 1973b.
- INEGE. Censo de Población 2015. República da Guine Ecuatorial. Resultados Preliminares.
- MINISTERIO DE ECONOMIA, P. E. I. P. Malabo: Dirección General de Estadística y Cuentas Nacionais, República da Guine Ecuatorial 2021.
- JOHNSON, D. C. *Language Policy*. London: Palgrave Macmillan, 2013.
- JOHNSON, D. C. Research Methods in Language Policy and Planning. In: TOLLEFSON, J. W. e PÉREZ-MILANS, M. (eds.). *The Oxford Handbook of Language Policy and Planning*. Oxford: Oxford University Press, 2018. p. 51-70.
- JUANOLA, J. *Primer paso á la lengua bubí, ó sea ensayo á una gramática de este idioma*. Madrid: A. Pérez Dubrill, 1890.
- KENYON, P. *Dictatorland: The Men Who Stole Africa*. London: Head of Zeus, 2018.
- KOILE, E.; GREENHILL, S. J.; BLASI, D. E.; BOUCKAERT, R. et al. Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences*, 119, n. 32, p. e2112853119, 2022. DOI: <https://doi.org/10.1073/pnas.2112853119>.
- KOUAM, G. *Le système nominal du banoo*. 1988. 115f. Dissertação (Mestrado em Línguas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé I, Yaoundé, 1988.
- KOUANKEM, C. *Complex Constructions in Bànò*. 2003. 110 f. Dissertação (Mestrado em Línguas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé I, Yaoundé, 2003.
- KRAUSS, M. The World's Languages in Crisis. *Language*, [s.l.], v. 68, p. 4-10, 1992. DOI: <https://doi.org/10.1353/lan.1992.0075>.
- LEMB, P. *Esquisse phonologique du mvumbò (ngumba)*. 1974. 100 f. Dissertação (Mestrado em Línguas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé, Yaoundé, 1974.
- LIPSKI, J. *The Spanish of Equatorial Guinea*. Tübingen: Max Niemeyer, 1985.
- LIPSKI, J. The role of the city in the formation of Spanish American dialect zones. *Arachne@Rutgers: Journal of Iberian and Latin American Literary and Cultural Studies*, New Brunswick, v. 2, n. 1, 2002a. DOI: <https://doi.org/10.14713/arachne.v2i1.21>.
- LIPSKI, J. The Spanish of Equatorial Guinea: research on la hispanidad's best-kept secret. *Afro-Hispanic Review*, [s.l.], v. 21, n. 1/2, p. 70-97, 2002b. DOI: <https://www.jstor.org/stable/20641705>.
- LIPSKI, J. The Spanish Language of Equatorial Guinea. *Arizona Journal of Hispanic Cultural Studies*, Phoenix, v. 8, p. 115-130, 2004. DOI: <https://doi.org/10.1353/hcs.2011.0376>.
- LIPSKI, J. *El español de América*. Madrid: Catedra, 2007.
- LIPSKI, J. Spanish-Based Creoles in the Caribbean. In: KOUWENBERG, S. e SINGLER, J. V. (ed.). *The Handbook of Pidgin and Creole Studies*. Oxford: Wiley-Blackwell, 2008. p. 543-564.

- LIPSKI, J. ¿Existe un Dialecto “Ecuatoguineano” del Español? *Revista Iberoamericana*, Pittsburgh, v. 80 p. 865-862, 2014.
- MACKEY, J. L. *A grammar of the Benga language*. New York: Mission House, 1855.
- MACKEY, J. L.; NASSAU, R. H. *Grammar of the Benga-Bantu Language*. New York: American Tract Society, 1892.
- MAGUGA, L. Gramática Kombe. In: *Diccionario español-kômbé*. Madrid: Instituto de Estudios Africanos, 1951. p. 15-88.
- MAHO, J. F. NUGL Online: The online version of the New Updated Guthrie List, Eberga referential classification of the Bantu languages. 2009.
- MANGADO, C. *Dialecto inglés-africano o broken-english de la Guinea Española*. Madrid: Estanislao Maestre, 1919.
- MASSAGA, M. W. *Le dialecte ngumba: essai descriptif*. 1971. 386 f. Tese de Doutorado, Faculté des lettres et sciences humaines, Université de Montpellier, 1971.
- NEGRÍN FAJARDO, O. *Historia de la Educación en Guinea Ecuatorial: el modelo educativo colonial español*. Madrid, UNED, 1993. 269 p.
- NERÍN, G. *Antropófagos, misioneros y guardias civiles*. Madrid: Los Libros de la Catarata, 2010a.
- NERÍN, G. *La última selva de España. Antropófagos, misioneros y guardias civiles. Crónica de la conquista de los fang de la Guinea Española, 1914-1930*. Madrid: Los Libros de la Catarata, 2010b.
- NERÍN, G. *Corisco y el estuario del Muni (1470-1931). Del aislamiento a la globalización y de la globalización a la marginación*. Paris: L'Harmattan, 2015.
- NISTAL ROSIQUE, G. El caso del español en Guinea Ecuatorial. In: *El Español en el Mundo: Anuario del Instituto Cervantes 2007*. Madrid: Instituto Cervantes, 2007. p. 73-76.
- OGOUAMBA, P. *Description phonétique et phonologique du kombé*. 1992. Dissertação Mestrado em Mémoire de DEA — Sciences du Langage, Université Lumière Lyon 2, 1992.
- ONDO-MÉBIAME, P. *Esquisse phonologique et morphologique du seki*. 1986. 148f. Dissertação (Mestrado em Linguística) — Faculdade de Letras e Ciências Humanas, Université Omar Bongo, 1986.
- PÉLISSIER, R. *Los territorios españoles de África*. Madrid: Consejo Superior de Investigaciones Científicas, 1964.
- PÉREZ, G.; SORINAS, L. *Gramatica de la lengua Benga*. Madrid: Editorial del Corazón de Mariá, 1928.
- POST, M. Fa d'Ambo. In: MICHAELIS, S. M.; MAURER, P., et al (ed.). *Portuguese-based, Spanish-based, and French-based Languages*. Oxford: Oxford University Press, 2013. p. 81-89.
- RAPONDA WALKER, A. Unpublished vocabularies of Sira, Kele, Okande, Njavi, Seke, Benga. 1916/1917.
- RÉNAUD, P. *Description phonologique et éléments du morphologie nominale d'une langue Pygmée du Sud-Cameroun: les Bajéle (Bipindi)*. Yaoundé, Direction Générale de la Recherche Scientifique et Technique (DGRST) Institut des Sciences Humaines, 1976.

- RURANGWA, I. M. Enquête linguistique sur le bubi, langue bantu insulaire de Guinée Equatoriale: phonologie et système des classes Rura. In: OBENGA, T. (ed.). *Les peuples bantou: migration, expansions et identité culturelle*. Libreville Paris: Centre International des Civilizations Bantu/L'Harmattan, 1987. p. 77-100.
- SALVADÓ Y COS, F. *Colección de apuntes preliminares sobre la lengua Benga: Ó sea, instrucción a una gramática de este idioma*. Madrid: A. Pérez Dubrull, 1891.
- SCHLUMPF, S. Hacia el Reconocimiento del Español de Guinea Ecuatorial. *Estudios de Lingüística del Español Actual*, Barcelona, v. 37 p. 217-233, 2016.
- SIROMA, J. L. *Esquisse linguistique du bapuku*. 1980. 69 f. Dissertação (Mestrado em Línguas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé I, Yaoundé, 1980.
- SKOLASTER, P. H. Die Ngumba-Sprache. *Mitt. des Seminars für orientalische Sprachen*, Berlin, v. 13, p. 73-132, 1910.
- SUNDIATA, I. K. Prelude to Scandal: Liberia and Fernando Po, 1880-1930. *Journal of African History*, London, v.15, n. 1, p. 97-112, 1974. DOI: <https://www.jstor.org/stable/180372>.
- SUNDIATA, I. K. *Brothers and Strangers*. Durnham: Duke UP, 2003.
- TOLLEFSON, J. W.; PÉREZ-MILANS, M. (ed.). *Oxford Handbook of Language Policy and Planning*. New York: Oxford University Press, 2018.
- TRILLES, H., 1935, Bruxelles. *Les Pygmées, leur langue et leur religion*. Imprimerie Médical et Scientifique, 1936.
- UM, E. N. *Morphologie Verbale du Mvùmbô*. 2002. 152f. Dissertação (Mestrado em Línguas e Linguística Africanas) — Faculdade de Artes, Letras e Ciências Humanas, Université de Yaoundé, Yaoundé, 2002.
- VAN HILLE, M. *Éléments de description du syntagme nominal en puku, langue bantue de zone A*. 1989. 190 f. Dissertação (Mestrado em Linguística) — Royal Museum for Central Africa, Université Libre de Bruxelles, Bruxelles.
- VILA, I. *Elementos de la gramática Ambú o de Annobón*. Madrid: A. Pérez Dubrull, 1891.
- WORLDBANK. Ahorro ajustado: gasto en educación (% del INB) - Equatorial Guinea. 2020.
- WORLDBANK. Equatorial Guinea. 2024.
- YAKPO, K. *A grammar of Pichi*. Berlin, Accra: Isimu Medi, 2009.
- YAKPO, K. *Gramática del Pichi*. Malabo: CEIBA Ediciones/Centros Culturales Españoles de Guinea Ecuatorial, 2010.
- YAKPO, K. Lenguas de Guinea Ecuatorial: de la documentación a la implementación. *Oráfrica*, Barcelona, v. 7, p. 13-28, 2011.
- YAKPO, K. Pichi. In: MICHAELIS, S. M.; MAURER, P., et al (ed.). *English-based and Dutch-based Languages*. Oxford: Oxford University Press, 2013. p. 194-205.
- YAKPO, K. *A grammar of Pichi*. Berlin: Language Science Press, 2019.
- ZAMORA SEGORBE, A. *Gramática Descriptiva del fa d'Ambô*. Madrid: 2010. 586 p.

Metaphorical Phraseologies in a Learner Corpus: Investigating Translations from Brazilian Portuguese to English

*Fraseologismos metafóricos em um corpus de aprendizes:
investigando traduções do português brasileiro para o inglês*

Jean Michel Pimentel Rocha
Instituto Federal de Mato Grosso do Sul
(IFMS) | Coxim | MS | BR
jean.rocha@ifms.edu.br
<https://orcid.org/0000-0001-8358-4774>

Abstract: This paper discusses the results of a study that aimed to examine whether Brazilian learners face difficulties in translating metaphorical phraseologies into English from a lexical, semantic, and cognitive perspective. The study drew on four main fields of research: metaphor, translation, phraseology, and corpus linguistics. We applied the conceptual metaphor theory and the cognitive-discourse perspective to identify the conceptual and linguistic metaphors in the corpus under study. Based on the cognitive translation hypothesis, mapping conditions were examined in the source and target languages. Additionally, grounded on corpus linguistics and phraseology, we evaluated the conventionality of the translated phraseologies in terms of their lexicogrammar, semantic, metaphor, and contextual features using a co-occurrence method and concordance line analysis. The research findings, along with related empirical studies, are presented and discussed. A total of 21 learners translated 81 phraseologies, resulting in 1202 translation solutions. The analysis showed that 86% of the translations represented metaphors with similar mapping conditions. However, 42% of them were understood as less or non-conventional metaphorical phraseologies. The findings suggest that the linguistic aspect posed more of a challenge in the translations rather than the conceptual one. Learners



retained the conceptual metaphor in the target language, but they often used unusual lexical choices in English due to the influence of their mother tongue's cognitive-linguistic system.

Keywords: metaphor; metaphorical phraseology; translation; learner corpus.

Resumo: Este artigo discute os resultados de um estudo que objetivou examinar se aprendizes brasileiros têm dificuldades ao traduzir fraseologismos metafóricos para o inglês, a partir de uma perspectiva léxico-sêmantica e cognitiva. A pesquisa se ancorou em quatro principais áreas: metáfora, tradução, fraseologia e linguística de corpus. A teoria da metáfora conceitual e a perspectiva cognitivo-discursiva foram empregadas para identificar as metáforas linguísticas e conceptuais no corpus de estudo. Com base na hipótese da tradução cognitiva, as condições de mapeamento foram examinadas na língua-fonte e na língua-alvo. Ademais, amparados pela linguística de corpus e pela fraseologia, a partir de análise estatística e de análise de linhas de concordância, a convencionalidade dos fraseologismos traduzidos foram avaliados em termos de aspectos lexicogramaticais, semânticos, metafóricos e contextuais. Os resultados da pesquisa são apresentados em articulação com pesquisas empíricas afins. Ao todo, 21 aprendizes traduziram 81 fraseologismos, o que resultou em 1202 esquemas de tradução. As análises mostraram que 86% das traduções representam metáforas com condição de mapeamento similar. No entanto, 42% delas foram classificadas como fraseologismos metafóricos pouco ou não convencionais. Os achados sugerem que o aspecto mais desafiador nas traduções foi linguístico e não conceptual. Os aprendizes mantiveram a metáfora conceptual na língua-alvo, mas em geral lançaram mão de unidades lexicais inco-
muns na língua inglesa sob influência do sistema linguístico-cognitivo de sua língua materna.

Palavras-chave: metáfora; fraseologismo metafórico; tradução; corpus de aprendiz.

1 Introduction

This paper presents the findings of a study (Rocha, 2020a) on how Brazilian learners translate metaphorical phraseologies from Portuguese to English, from a lexical, semantic, and cognitive perspective. First, it was examined whether learners' phraseometaphorical choices were conventional in the English language, considering lexicogrammar, semantic, metaphor, and contextual features. Additionally, it evaluated whether the conceptual mappings underlying the phraseologies were similar or different when comparing the source language (SL) and the target language (TL).

Regarding the lexicogrammar and the semantic-metaphorical nature, the study aimed to determine if English learners use conventional or unconventional metaphorical phraseologies in their writing. The basic premise of this research is that phraseologies play a crucial role in enhancing foreign languages learners' lexical competence. According to Sinclair's idiom principle (1991), a language speaker has a collection of ready-made expressions in their mental lexicon that can be used to express coherent and natural utterances during communication. These expressions, which are known as phraseologies, are linguistic repetitions that have been passed down from one generation to the next and represent the conventional nature of language on different levels – lexical, syntactic, semantic, and pragmatic (Tagnin, 2013).

Learners of a foreign language may not always have access to ready-made expressions in their mental lexicon that they can immediately use. Not being aware of what is socially shared may cause them to create uncommon linguistic constructions that are scarcely or not at all observed in a linguistic community. When communicating, as pointed out by Alves and Tagnin (2010), they may be perceived as an innocent speaker/hearer (Fillmore, 1979), someone who is not aware of what is institutionalized in a language and therefore relies on literal or atypical usage. The idea of an innocent speaker/hearer can be reinterpreted from the perspective of a translator who is still learning and is not familiar with the conventions of the languages they are working with. This translator believes that there is complete equivalence between the languages and that a literal translation is possible or that each lexical item of the source text must be translated (Alves; Tagnin, 2010).

Considering the cognitive aspect, this work also relies on the guiding assumption that metaphors don't simply conceptualize a pre-existing reality, but rather create a new one (Lakoff; Johnson, 2003 [1980]). When it comes to translation, learners must face the challenge of dealing with a complex reality that can be categorized linguistically, metaphorically, and culturally. This can lead to changes in both the linguistic and conceptual levels. Therefore, "to translate a metaphor means to stand face to face with the kind of reality that the metaphor has construed" (Arduini, 2014, p. 41).

Combining both the linguistic and the conceptual viewpoints, we attempted to answer the following questions. Firstly, if phraseologies can be difficult for translation learners, then how would they react to the metaphors, which may be part of the translating process? In other words, would there be difficulties in translating phraseologies due to the underlying conceptual mappings? Secondly, regarding lexical choices, how could the phraseometaphor-

ical solutions proposed by the learners be characterized? Would they be conventional in the English language taking into account the lexicogrammar, semantic, and metaphorical contexts in which they are embedded?

To answer these questions, the study relies on four main areas: metaphor, translation, phraseology, and corpus linguistics. The first section provides information about the study corpus and explains how these fields are used to address the outlined questions. The second section presents research findings, comparing them with similar empirical studies that reflect on lexicon and metaphor from a pedagogical perspective: Boers (2000), Charteris-Black (2002), Danesi (1994, 2008), Deignan, Gabrys and Solska (1997), Irujo (1986), Littlemore and Low (2006a, 2006b), and Philip (2010). It also considers studies that focus on metaphor translation from a cognitive approach: Baiocco and Siqueira (2018), Kovécses (2014), Mandelblit (1995), Martins (2008) and Rodríguez-Marquez (2010). The final section examines the implications of the findings for foreign languages teaching, with a particular focus on translation teaching and learning.

2 Theoretical and methodological background

A corpus of learner translations was compiled to analyze the phenomena under investigation. This corpus is a part of the Multilingual Student Translation (MUST) project (Granger; Lefer, 2018), which aims to develop a multilingual translation corpus to aid research in learner corpus and translation studies. The Brazilian component¹, which is coordinated by Prof. Dr. Adriane Orenha-Ottaiano at São Paulo State University (UNESP), has the following characteristics: (i) it consists of 20 newspaper articles (500-800 words) written in contemporary Brazilian Portuguese. The articles cover a variety of subjects such as terrorism, politics, LGBT activism, environment, education, and arts; (ii) it is a parallel corpus translated from Portuguese to English; (iii) the maximum number of participants involved in this project is 21, who are undergraduate students pursuing a bachelor's degree in Languages (Translation) and a licentiate degree in Languages. These students have varying levels of proficiency, ranging from intermediate to advanced. The study corpus is divided into two sub-corpora – the Portuguese sub-corpus and the English sub-corpus. The Portuguese sub-corpus contains 13,520 tokens, while the English sub-corpus contains 205,176 tokens, both computed in the Sketch Engine platform. These corpora were used as the basis for extracting metaphorical phraseologies.

2.1 Metaphor studies

This research draws upon two theoretical frameworks in the metaphor studies literature: the conceptual metaphor theory and the cognitive-discourse perspective. The conceptual metaphor theory views metaphor as a relational mechanism that involves experiencing one thing in terms of another. This theory was used to determine the source and target domains of the conceptual metaphors underlying the phraseologies analyzed (Lakoff; Johnson, 2003 [1980]).

¹ The Research Ethics Committee of IBILCE/Unesp approved the data collection under the number 94053718.4.0000.5466.

To experience something in terms of another is to understand a conceptual domain – “any coherent organization of experience” (Kövecses, 2010, p. 4) –, defined by another one. According to Lakoff (1987), conceptual domains are a result of Idealized Cognitive Models (ICM) or gestalt experiences, a coherent whole structured on human natural experiences. They are products of human nature (Lakoff, 1987; Lakoff; Johnson, 2003 [1980]) in interaction with the body (implying its perceptual-motor, mental and emotional dimensions), with the physical environment (space-motion, object manipulation, food habits, etc.), and with other people culturally situated (in social, political, economic, and religious levels).

Gestalt experiences, because human activity in the world, triggers the creation of metaphorical domains: the source domain (SD) that structures/defines the target domain (TD). The SD is a more concrete conceptual domain (e.g. physical orientation, objects, substances, plants, money, vision, travel, war, food, construction, games) that is derived, for instance, from linguistic evidence. The TD, in turn, is characterized by being a more abstract notion (e.g. love, time, ideas, people, understanding, argument, happiness, health, status, politics) that assists in the conceptualization of a SD (Lakoff; Johnson, 2003 [1980]; Kövecses, 2010).

Based on the conceptual metaphor theory, especially in Lakoff and Johnson (2003 [1980] and Kövecses (2005, 2010), the conceptual metaphors underlying the phraseologies raised in the corpus were identified. To guide this step, we consulted the Master Metaphor List (Lakoff; Espenson; Goldberg, 1989) and the MetaNet Metaphor Wiki².

The metaphorical phraseologies in Portuguese were categorized into different domains for further interlanguage analysis. Table 1 provides an illustration of these domains, along with examples extracted from the Portuguese study corpus, highlighting the metaphorical phraseologies and their respective conceptual metaphors.

Table 1 – Source domains identified in the study corpus in Portuguese

Domains	Examples from the study corpus in Portuguese	Conceptual metaphor
Container	<i>Temer segue de pé e impermeável às críticas daqueles que questionam sua legitimidade e a dureza de seus ajustes.</i>	THE BODY IS A CONTAINER
Motion/Direction	<i>Nós seguimos pelo caminho errado”, disse... ao anunciar as linhas que seu governo seguirá.</i>	MEANS ARE PATHS TO A PURPOSE
Control	<i>Seu governo está na corda bamba desde que o jornal O Globo revelou...uma comprometedora gravação de uma conversa com o empresário.</i>	CONTROL IS UP/LESS OF CONTROL IS DOWN
Plant	<i>A edição das medidas provisórias que reduzem áreas protegidas...está na raiz desse problema.</i>	PROBLEM IS A PLANT
Construction	<i>Meus argumentos estavam alicerçados em literatura e publicações científicas.</i>	ARGUMENT IS CONSTRUCTION
War/physical conflict	<i>Temer conseguiu ganhar tempo e sobreviveu ao julgamento... ainda há muitas frontes abertas.</i>	POLITIC IS WAR/PHYSICAL CONFLICT

² Repertoire MetaNet Metaphor Wiki: https://metaphor.icsi.berkeley.edu/pub/en/index.php/MetaNet_Metaphor_Wiki. Access in 2019/2020.

Games/sports	<i>A frase é do presidente americano...e joga no nosso colo a enorme responsabilidade de virar o jogo.</i>	POLITICAL ACTIVITIES ARE GAMES/SPORTS
Money/ economic transactions	<i>Os anos de abuso... de recursos naturais... vinham matando rios.... A natureza cobrou seu preço.</i>	NATURE EXPLOITATION IS AN ECONOMIC TRANSACTION
Play	<i>O ex-ministro da Casa Civil José Dirceu tem evitado os holofotes.</i>	POLITICAL ACTIVITY IS A PLAY
Other domains	<i>Muitas levaram cantadas machistas e engolirão a seco o vazio do respeito que não existe no dia a dia de muitas dessas profissionais.</i>	ACCEPTING IS SWALLOWING

Source: created by the author.

Before identifying the conceptual mappings in the study corpus in Portuguese, we identified the linguistic metaphors. We initially selected possible candidates for metaphorical phraseologies and then observed if they had an abstract reference that was semantically related to a distinct use of its concrete meaning. To improve the methodological strategy, we used a linguistic procedure developed under a cognitive-discourse approach to metaphor: the Metaphor Identification Procedure (MIP), created by the Pragglejaz Group (2007). This procedure intends to investigate the communicative functions of metaphor in text/discourse (Steen, 2007; Verezza, 2016). To establish the metaphorical meaning, they assumed the omnipresence of metaphor in language, aiming to explore its usage in different genres. Once the usage has been set as a starting point, the Group defines metaphorical meanings as the ones that emerge from the contrast between the contextual meaning of a lexical unit and its basic meaning. The basic meaning is absent from the context under investigation but is observed in other contexts (Steen et al., 2010). The phraseology highlighted below exemplifies the contrast between the basic and contextual meaning:

*Em leitura e em matemática, que são as outras duas áreas avaliadas pelo Pisa, o quadro também é desalentador: seguimos no grupo dos lanternas, com a 59^a e 66^a colocações, respectivamente. [...] Todo esse cenário mostra o **fundo do poço** a que chegamos, algo que as próprias avaliações brasileiras já vinham, ano a ano, indicando.*

In this excerpt, the phraseology *fundo do poço* has a contextual meaning that contrasts with and can be understood in terms of its basic meaning. Contextually, it means a difficult situation and is used to describe Brazil's poor performance in an educational assessment. The basic meaning of the phraseology is a hole dug to reach underground water. In the educational scenario, being in the *fundo do poço* is similar to hitting the bottom of a well and struggling to find a way out.

Through the MIP procedure, we were able to classify approximately 200 phraseologies as metaphorical. However, only 81 of them were considered for further cross-linguistic analysis as they met the statistical threshold in tests such as T-score, MI, and LogDice, and were predominantly metaphorical in the concordance lines analysis based on the Portuguese reference corpus (Portuguese Web 2011 - ptTenTen). The research only considered conventional metaphorical phraseologies in Portuguese and did not include weaker associations and discourse-specific metaphors (Steen, 2007; Verezza, 2013), which are not institutionalized in the language.

To summarize it, the conceptual metaphor theoretical and methodological framework, along with the cognitive-discourse perspective, supported the identification of linguistic and conceptual metaphors in the Portuguese study corpus. The mapping correspondences were described in terms of SD and TD, and the conceptual metaphors were analyzed, enabling the analysis of Portuguese phraseologies and their translations to English. This comparison helped determine whether the mapping conditions were similar or different between the two languages.

2.2 Metaphor and translation

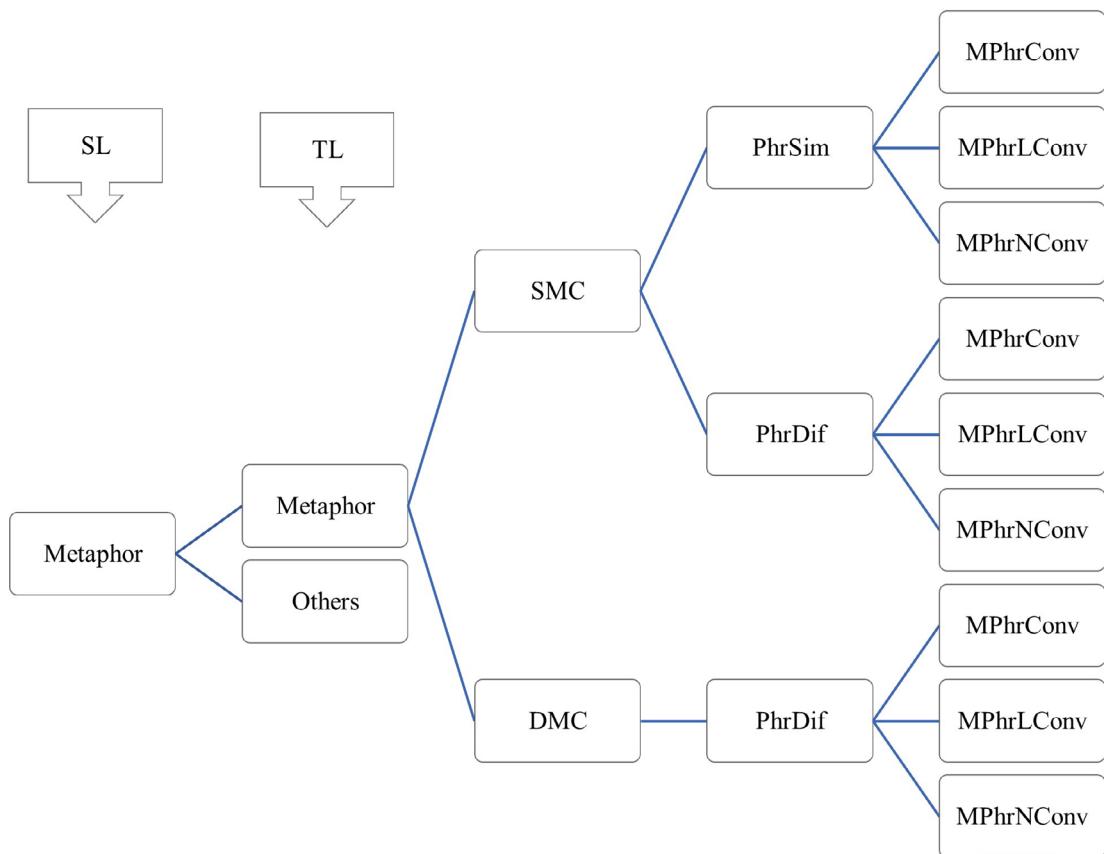
The reality constructed via metaphor can involve different conceptual systems. The difference among them, according to researchers such as Mandelblit (1995) and Kövecses (2014), would explain difficulties in translation. Mandelblit's cognitive translation hypothesis, proposed in 1995, argues that difficulties in translation arise due to a lack of correlation between the conceptual systems of the source language (SL) and the target language (TL). In other words, the process of translation involves transferring information from one system to another. This transfer is faster when there is a similar mapping condition (SMC) between the two languages, compared to a different mapping condition (DMC), where the changes would be both linguistic and conceptual. According to Mandelblit (1995), the more cultural aspects shared by the languages, the higher are the chances of a similar mapping condition.

In Mandelblit's view, a translation that requires a conceptual system transfer can be seen as a problem-solving process. The translator aims to find a linguistic expression in the target language (TL) that conveys the same communicative purpose as the one in the source language (SL). During the process, they may experience a temporary functional fixation³, which is a state where they remain focused on the SL metaphorical system and struggle to find an equivalent expression in the TL.

By emphasizing the cognitive aspect of metaphors and proposing a hypothesis for how the translation process occurs, Mandelblit's work has encouraged cognitive-based translation research (e.g. Al-Hasnawi, 2007; Baiocco; Siqueira, 2018; Maalej, 2008; Martins, 2008; Rodríguez-Marquez, 2010; Taheri-Aradali, Bagheri, Eidy, 2013). Unlike Mandelblit (1995), the focus of these researchers was not on the process, but on the product of translation. To investigate it, they mobilized possible translation schemes, which involved, for example, the comparison of mapping conditions, linguistic realization, among other strategies. Our research follows a similar approach to Taheri-Ardali, Bagheri, and Eidy's work (2013), designing a scheme to test the hypothesis that focuses on translation as a product. The reason behind creating a new schema was due to the nature of our data, which resulted from translations done by learners. Since no translation equivalence was established previously, there were several translation possibilities available. Our scheme includes the conventional aspect of metaphorical phraseologies as well.

³ According to Mandelblit (1995) this idea has its origins in the Gestalt School, which perceives the problem resolution as a process in which who tries to solve it do so based on insights from its structure. Failures in its resolution would be explained by functional fixation, to the focus given to the original formation of the problem, which would impede its reconstruction and solution.

Figure 1 – Translation schemes of metaphorical phraseologies



Source: adapted from Rocha (2020a, p. 88) based on Taheri-Ardali, Bagheri, Eidy (2013).

The abbreviation **SMC** (Similar Mapping Condition) refers to the phraseology in the source language (SL) that shares a similar conceptual mapping to the phraseology of the target language (TL). This similarity is based on the source and target domains, for example, *barril de pólvora* and *powder keg* share the same conceptual metaphor: ANGRY PERSON IS A CONTAINER ABOUT TO EXPLODE, as mapped in the SL. On the other hand, **DMC** (Different Mapping Condition) refers to the phraseology in the SL that does not share the same conceptual mapping to the phraseology of the TL. For instance, *trazer à tona* and *bring to the table* have different mappings. In *trazer à tona*, one can distinguish an image of an object submerged in a container, which is not observed in *bring to the table*.

PhrSim and **PhrDif** refer to phraseologies with similar and different lexical implementations, respectively. This relation is quite complex to establish. When studying different languages, Kövecses understands that both have unique lexical implementations. Taheri-Ardali, Bagheri, and Eidy (2013) reported difficulties in defining it, but they acknowledge the distinction proposed by Mandelblit (1995). According to Mandelblit, in the case of similar lexical implementation, there would be a word-to-word translation.

In this paper, the distinction made by Mandelblit was maintained. The lexical form of the Portuguese phraseology was compared with the translations made by the learners. During this process, cognate forms and literal or *prima facie* translations were also taken into consideration. For instance, the phraseology *barril de pólvora*, when compared to *barrel of gunpowder*, was distinguished as PhrSim, whereas in comparison to *powder keg*, it was recognized

as PhrDif. In this case, the correlation at the word level was established through the use of cognates and the syntagmatic structure. If a learner altered the sequence of words within a phraseology, while using the same words (e.g., changing *barrel of gunpowder* to *gunpowder barrel*), the phraseology was still identified as PhrSim. Before parasyntactic forms, there were certain difficulties in translating phraseologies such as *debaixo do tapete* which can be translated as *under the carpet*, *under the mat*, or *under the rug*. These translations were all recognized as PhrSim since they are possible direct translations. At this analytical stage, although the semantic aspect cannot be completely disregarded, the lexical form took precedence.

In our analysis of the phraseology's conventional aspects, we approached the subject from a lexicogrammar and semantic-metaphorical perspective, using the conventionality levels proposed by Tagnin (2013). Therefore, from a lexicogrammar point of view, we examined the associative nature of phraseologies by conducting statistical tests which helped us confirm if the learner's choices were statistically significant and not random combinations. From a semantic-metaphorical perspective, the metaphorical meaning of translated phraseologies in English were interpreted by examining concordance lines in a reference corpus. The analysis also took into account the contextual meaning in both languages. As it focused on translations by learners, three classes were differentiated: **MPhrConv**, **MPhrLConv**, and **MPhrNConv**.

- ◆ **MPhrConv (Conventional Metaphorical Phraseology):** it refers to the identification of statistically significant phraseologies in two or more measures ($T\text{-score} > 2$, $MI > 3$, and $LogDice > 3$). These phraseologies should exhibit a predominance of metaphorical occurrences that are semantically compatible with the phraseology of the source language. This is determined by analyzing the concordance line samples in the English reference corpus.
- ◆ **MPhrLConv (Less Conventional Metaphorical Phraseology):** it refers to two types of phraseologies. The first type includes statistically relevant phraseologies that have less metaphorical occurrences and more basic/non-metaphorical meanings. The second type includes phraseologies with statistical measures below the proposed values, but with a predominance of metaphorical occurrences that are semantically compatible with the phraseology of the SL.
- ◆ **MPhrNConv (Non-Conventional Metaphorical Phraseology):** it refers to phraseologies, above or below the statistical measures, semantically incompatible with the context of the expression in the SL; it also includes occurrences that did not appear in the list of collocates of a node. In this case, as well as in the others, the conservation of the wording *metaphorical phraseologies* is due to the learner's own construction that is metaphorical in the translations when considered a contrast between contextual and basic meanings.

The category **Others** was applied when the learner did not use a metaphor, but instead used a non-figurative lexicon, paraphrased the meaning, or omitted the lexical unit that triggered the metaphor. This category also included instances in which the learner did not resort to a phraseology (whether it was metaphorical or not) or employed a simple word or compound.

The adoption of the mentioned translation schemas was fundamental to achieving the research goals. Their application to the data enabled us to identify the phraseologies mapping conditions and their conventional aspects, providing an overall picture of learners' translation trends.

2.3 Metaphor, phraseology, and corpus linguistics

The lexicon is a special place to explore metaphors, and one way to do that is through phraseologies. They are linguistic representations of a social group's collective identity. Saussure's (2012 [1916]) concept of language as a social institution, as well as Coseriu's (1979) distinction of language system and norm, as stated by Zuluaga, are significant to understand the linguistic status of phraseologies. According to Zuluaga (1980, p. 21), they represent "the collective heritage of a linguistic community, [...] part of the repertoire of linguistic elements that precedes the discourse and are known by its speakers. They are normal traditional elements repeated in the linguistic activity"⁴. Their reproduction in chunks, as part of the idiom principle (Sinclair, 1991), reinforces their importance for foreigners or second language learners.

Since phraseologies are shared by a social community and are repeated in lexical bundles, in different nuances, they have fixation as a distinctive feature, not only in the system level, considering semantic and syntactic aspects, but also in the norm level, which considers language usage (Zuluaga, 1980). According to the theoretical views of Ruiz Gurillo (1997) and Corpas Pastor (1996), which align with Zuluaga's approach, phraseologies have the following properties:

- 1 They are pluriverbal units, which means they consist of at least two words.
- 2 They are institutionalized, meaning their use is conventional and recognized by the speakers of a particular language, and they occur frequently in discourse.
- 3 They are established in syntagmatic and semantic terms.
- 4 They are subject to variation due to sociolinguistic and contextual factors.
- 5 They may or may not be idiomatic or compositional, meaning the unit's meaning may or may not be the sum of its parts. They can also result from metaphorical processes.
- 6 These properties exist on a continuum rather than being categorical.

The properties mentioned above include a range of phraseologies that are difficult to identify precisely. According to Ruiz Gurillo (1997, p. 85), they are units "of fuzzy borders, of malleable and movable limits"⁵. In this research, aware of the tenuous boundary that delimits phraseologies, we acknowledge that they are non-discrete units, following the works of Corpas Pastor (1996), Ruiz Gurillo (1997), Philip (2011), and Rocha (2017). Acknowledging it does not necessarily mean denying the possibility of distinguishing them, but in some cases, it can be challenging due to overlapping criteria. Additionally, the criteria used may be influenced by the speaker's subjective evaluation, which goes beyond the linguistic level (Philip, 2011).

⁴ [...] al patrimonio colectivo de la comunidad lingüística, forman parte del acervo [...] o repertorio de elementos lingüísticos, anteriores al hablar, conocidos por los hablantes. Son elementos tradicionales 'normales' y repetidos en la actividad lingüística.

⁵ [...] fronteras difusas, de límites maleables y móviles.

Making a clear distinction between different types of phraseologies is beyond the scope of this research. Thus, our focus is on referential phraseologies, which convey information related to daily life and world entities (Granger; Paquot, 2008) without being linked to a specific category. Metaphor is the key element that connects them, and for this reason, in this paper they are being referred to as *metaphorical phraseologies*: a pluriverbal unit, statistically relevant, sanctioned and fixed by the linguistic norm of a society. It is used in varied socio-communicative contexts, formally structured in a syntagm, and semantically featured by (relative) compositionality/idiomaticity, as well as for its metaphoricity.

This perspective of understanding metaphorical phraseologies favors a dialogue between two major branches in phraseological studies: the phraseological approach, which concentrates on linguistic properties; and the frequency-based approach or distributional approach (Evert, 2005; Nesselhauf, 2005; Granger, Paquot, 2008), which highlights the frequency of co-occurrence of lexical units.

The frequency-based approach is closely associated with corpus linguistics since it originated within its scope. Along with other areas, corpus linguistics plays a significant role in the development of this research. By considering language as a probabilistic system, corpus linguistics reveals what is likely to happen in language usage, highlighting the systematic nature of linguistic patterns (Berber Sardinha, 2004; Halliday, 1991). Consequently, it rejects the notion of randomness in language and reinforces the idea that language is highly organized and planned (Sinclair, 1996). This organization can also be observed in metaphorical terms, either on the linguistic or on the conceptual level.

The use of a corpus can be a valuable tool to understand language and the cognitive concepts it represents. A corpus linguistics approach, in its essence inductive (Tognini-Bonelli, 2001), has the potential to redefine and reorient linguistic theories. It can also complement existing research that focuses on investigating metaphor. Examples of such research include studies by Deignan (1997, 2005), Berber Sardinha (2011a, 2016), Semino (2008), and Rocha (2020a, 2020b).

Corpus linguistics applied to metaphor studies represents a cognitive-discourse shift that uses language as a data source to understand metaphor in thought (Vereza, 2013). Deignan's (1997) works are a notable example of this approach. Deignan (1997) argues that while the metaphor conceptual theory can apprehend the linguistic metaphor nature, it is not complete without the use of corpus. By examining syntactic, collocational, and semantic patterns, which were not prioritized in previous investigations due to the limitations of intuition, the author reinforces the ubiquity of the phenomenon in language and its conventionality.

In this study, the analysis of conventional metaphorical phraseologies is based on linguistic evidence and patterns identified using the epistemological perspective of corpus linguistics, in conjunction with the frequency-based approach (Granger, Paquot, 2008). The procedure used is based on the co-occurrence of patterns (nodule + collocate) that are statistically determined, following the work of Sinclair, Jones and Daley (2003 [1970]) and Berber Sardinha (2011a, 2012). The platform Sketch Engine (Kilgarriff et al., 2014) was used to access the reference corpora Portuguese Web 2011 – ptTenTen and English Web 2015 – enTenTen, which assisted in comparing the learner's phraseometaphorical choices.

The Sketch Engine platform was useful for two methodological steps. The first step aimed to confirm the *node + collocate* co-occurrence was not random, by using a reference corpus in Portuguese and measures such as T-score, MI (Mutual Information), and LogDice. The

researcher initially selected phraseologies based on intuition and followed the MIP application. In the next stage, a statistical threshold was set in the Sketch Engine, which required a phraseology to have values of $MI > 3$, $T\text{-score} > 2$ and $\text{LogDice} > 2$ to be included in the comparative analysis between learner's translations and reference corpora. Like Berber Sardinha's (2011a) decision, a phraseology needed to exceed at least two of these values to be considered for the analysis.

The second step in the platform consisted of a comparative analysis which involved using concordance tools and statistical measures to determine if the learner's phraseometaphorical choices were conventional in the TL context. These measures are widely used in corpora research and are generated by the platform. Overall, they show the strength and exclusivity of the associations. According to the Sketch Engine Platform, the $T\text{-score}$ measure shows that a word association is not random, indicating the certainty of the association. The Mutual Information score measures how often words co-occur compared to how often they appear separately. The LogDice measure indicates the typicality or strength of a co-occurrence. The values applied in this research are based on similar studies conducted by Berber Sardinha (2011a), Brezina, McEnery and Wattam (2015), Durrant and Schmitt (2009), Frankenberg-Garcia (2018), McEnery (2006), and Stubbs (2001).

It's worth noting that the use of metaphors was another factor considered when evaluating the phraseologies in the comparative analysis. In addition to passing tests, it was necessary to check whether they were mostly used in a metaphorical context in the concordance lines. However, examining them when the occurrences are high is not feasible. To address this issue, a statistically random sample of the total concordance lines was used, obtained through the "get a random sample" command of the Sketch Engine platform. For both languages, we analyzed one hundred lines of phraseologies with over one hundred occurrences. Our goal was to determine if the phraseologies were used in a metaphorical context regardless of the conceptual metaphor of the source language. Therefore, phraseologies that exceeded the threshold but did not show a predominance of metaphorical occurrences were discarded.

3 Research findings

As previously mentioned, 81 phraseologies were manually identified in 20 texts in Portuguese. They resulted in 1202 proposed translations from a maximum of 21 participants. In this section, we present some considerations based on the translation schemes applied by the students. Table 2 shows a quantitative distribution of the schemes by SD. The number in parentheses after each domain refers to the total numbers of phraseologies mapped in Portuguese for that domain:

Table 2—General distribution of translation schemes

Translation schemes	Total number of schemes translated by domain											Total/%
	Container (9)	Motion/Direction (13)	Control (9)	Plant (6)	Construction (5)	War/physical conflict (12)	Games/sports (4)	Money/Economic trans. (4)	Play (3)	Other domains (17)		
SMC/PhrSim/MPhrConv	22	67	60	28	3	47	10	0	15	51	303/1202 (25%)	
SMC/PhrSim/MPhrLConv	14	13	8	5	6	13	3	16	8	26	112/1202 (9%)	
SMC/PhrSim/MPhrNConv	4	18	0	15	10	3	0	0	1	38	89/1202 (7%)	
SMC/PhrDif/MPhrConv	49	24	25	11	28	49	25	9	3	50	273/1202 (23%)	
SMC/PhrDif/MPhrLConv	22	18	18	3	14	21	10	20	2	16	144/1202 (12%)	
SMC/PhrDif/MPhrNConv	6	27	10	7	10	14	2	12	2	35	125/1202 (10%)	
DMC/PhrDif/MPhrConv	3	1	2	2	0	4	9	0	0	2	23/1202 (2%)	
DMC/PhrDif/MPhrLConv	0	1	0	6	2	0	0	0	0	2	11/1202 (1%)	
DMC/PhrDif/MPhrNConv	4	7	1	0	1	5	0	0	0	11	29/1202 (3%)	
Others	7	9	7	5	1	12	4	1	15	32	93/1202 (8%)	
Total	131	185	131	82	75	168	63	58	46	263	1202/100%	

Source: adapted from Rocha (2020a, p. 169)

Looking at the Table, we can see that the *others* category accounts for 8% of all schemes. The data analyzed includes non-metaphorical occurrences, such as paraphrases of the source language's wording, as well as omissions and instances in which a phraseology was not used. This can be seen through examples (1) to (3):

- (1) *Muitas levarão cantadas machistas e engolirão a seco o vazio do respeito que não existe no dia a dia de muitas dessas profissionais.*

Many gonna rather **accept** and reproduce the sexist, homophobic and racist jokes.

- (2) *O que vemos no noticiário político ou nas páginas policiais – cada vez mais semelhantes entre si – é mais um ato do que se desenrola nesse mesmo pano de fundo.*

What we see in the political news or in police pages - each time more alike - is one more act of what is going on at this same **background**.

- (3) *A desconfiança passou a orientar a ação dos legisladores governistas, que aprovaram sanções à Rússia contra a vontade da Casa Branca e deram passos para impedir que o presidente demita seu secretário de Justiça...*

The suspicion began to guide the actions of government legislators, which have approved sanctions to Russia against the will of the White House and **O** prevented that the President dismisses his...

In (1), one of the learners suggested the paraphrase *accept* to replace the phraseology *engolir a seco*. The learner probably used this strategy because they were not familiar with a conventional metaphorical phraseology in the target language. For instance, *swallow hard* is a phraseology that expresses the acceptance of an idea grudgingly. It appears 2,906 times in the reference corpus in English and has a T-score of 53.32, MI of 6.52 and LogDice of 4.69. In this case, it is acceptable to use a certain interpretation of the SL phraseology, even though it might not be entirely appropriate in the context. However, this approach can result in a loss of the metaphorical imagery.

In example (2), the phraseology *pano de fundo* was replaced by a simple word *background*, but the mapping condition of the SL was preserved by the conceptual metaphor **LIFE IS A PLAY**. In example (3), the learner chose to omit the SL phraseology, which is based on the conceptual metaphor **MEANS ARE PATHS TO A PURPOSE**. Alternatively, the learner could have used the metaphorical phraseology *take a step*, which has SMC and is statistically significant (74,640 occurrences, T-score 254.41, MI 3.86 and LogDice 6.51).

Metaphors, particularly linguistic ones, serve a textual function by highlighting an idea, establishing relationships, or creating some communicative effect in the reader. However, these functions are not always conveyed in translations when literal solutions are proposed, leading to a loss of semantic features. Nevertheless, in our data, this loss is not significant, as 92% of learners attempted to maintain metaphorical phraseologies in both languages, using either similar or different mapping conditions (SMC/DMC).

It was observed that 86% of the translations contain SMC, regardless of whether they are similar or different in nature, and whether they are conventional, less conventional, or non-conventional. This high percentage indicates that most learners have retained the conceptual metaphor of the source language in the target language. This can be attributed to their L1's cognitive, linguistic, and cultural interference. However, this percentage has a double implication that needs to be considered.

The first one is favorable to the learner. It would be related to the fact that in the phraseologies analyzed, both languages share, to an extent, cognitive, linguistic, and cultural traditions that make them similar in terms of mapping conditions. This could lead to conventional translation solutions when the learner uses their knowledge of their mother tongue's lexical, semantic, and conceptual structures. Besides that, human action in the social and physical environment – the embodiment role (Gibbs, 2006; Kövecses, 2010; Lakoff, 1987) – plays a significant role in the mapping maintenance. The domains presented in this research

are perceptible in both Portuguese and English languages, indicating that they share common ways of categorizing reality.

Linguistically, SMC is evidenced by similar lexical realization (PhrSim) through *prima facie* translation, examples (4) and (5), which are linked to the conceptual metaphors PROBLEM IS PLANT and POLITIC IS WAR/PHYSICAL CONFLICT.

- (1) *A edição das medidas provisórias que reduzem áreas protegidas e regularizam a grilagem está na raiz desse problema.*

The change of the provisional measures that reduce protected areas and regularize the land grabbing is in the **root of this problem**.

- (2) *O novo líder francês vê essa posição, que já vinha sendo defendida... como uma oportunidade para reforçar a unidade entre os 27 países do bloco.*

The new French leader sees this **position**, which has already been **defended** by the British Prime Minister Theresa May's direct advisers, as an opportunity to reinforce the unity among the 27 countries that constitute the bloc.

Also, SMC is expressed using different lexical realization (PhrDif) and conceptual metaphors, such as ACCEPTING IS SWALLOWING and INSISTENCE IS CONTINUOUS REPETITION, as exemplified in (6) and (7).

- (3) *Muitas levarão cantadas machistas e engolirão a seco o vazio do respeito que não existe no dia a dia de muitas dessas profissionais.*

Many will be chatted up with in a sexist way and **swallow hard** the empty space of respect, which doesn't exist in the daily life of those professionals.

- (4) *Um bom trabalho diplomático com os vizinhos poderia ter melhor efeito do que deixar nossas polícias enxugando gelo.*

A good diplomatic work with the neighbors could have a better effect than leaving the **police herding cats**.

In nearly half of the cases (48%), translations were made without any significant challenges due to the overlapping linguistic and conceptual features among the phraseologies, using SMC and MPhrConv. When learning a new language, it's common for the learner to rely on their mother tongue as a starting point. This is because the learner's native language is their main point of reference, especially when faced with expressions that require translation. Therefore, embracing the source language might be an effective strategy for a successful translation.

Nesselhauf (2005) studied the production of lexical patterns used by learners in German and English. Although she did not assess the role of metaphor, she pointed out the beneficial influence of the mother tongue. The cognitive-linguistic proximity between the mother tongue and the target language can be an advantage, as observed by Boers (2000), Charteris-Black (2002), Danesi (1994, 2008), Deignan, Gabrys, and Solska (1997), Littlemore et al. (2014) and Irujo (1986). However, it can also have the opposite effect, leading the learner to make uncommon choices. This brings us to the second implication: the negative interference of the mother tongue. It was noted that some learners stick to Portuguese, maintain SMC, but use less conventional phrases in the analyzed contexts (MPhraLConv or MPhraNConv - 38%), as can be seen in examples (8) and (9):

- (5) *Temer segue de pé e impermeável às críticas* daqueles que questionam sua legitimidade e a dureza de seus ajustes.

But Temer follows up and **impermeable to criticism** from those who question his legitimacy and the toughness of his adjustments.

- (6) *Parece que somos um barril de pólvora* a procura de uma faísca.

It seems that we are a **barrel of gunpowder** looking for a spark.

In (8), *impermeável às críticas* is the linguistic realization of the metaphor THE BODY IS A CONTAINER (Lakoff; Johnson, 2003). In this example, Michel Temer's body is compared to a closed container, which is impervious to any substances, in this case, negative criticism that could harm his political reputation. One of the learners suggested *impermeable* as a direct translation, but it was not found in the co-occurrence list of *criticism*. From a cognitive-linguistic perspective, the metaphor is preserved, but *impermeable to criticism* is not a conventional metaphorical phraseology. Alternative phraseologies that are statistically relevant and appropriate to the context include *impervious to criticism* (72 occurrences, T-score: 8.44 MI: 7.57 and LogDice: 2.51) and *immune to criticism* (371 occurrences, T-score: 18.89 MI: 5.70 and LogDice: 4.06).

Another phraseology which has THE BODY IS A CONTAINER as an underlying conceptual metaphor is *barril de pólvora* (example 9). Specifically, the underlying metaphor is ANGRY PERSON IS A CONTAINER ABOUT TO EXPLODE, resulting from the metaphor association THE BODY IS A CONTAINER FOR THE EMOTIONS, EMOTION IS HEAT AND ANGER IS HEAT (Kövecses, 2010). The phraseology *barrel of gunpowder* (208 occurrences, T-score: 14.41 MI: 10.44 and LogDice: 4.67), although statistically significant, is MPhrLConv because of the basic meanings predominance in the concordance lines examined. However, this translation is not entirely accurate because it mainly refers to a wooden container filled with gunpowder (example 10). A more suitable translation (MPhrConv) would be *powder keg* (2,332 occurrences, T-score: 48.29 MI: 14.02 and LogDice: 8.60), which is a metaphorical term used to describe a situation that is likely to become dangerous or explosive, as suggested in (11):

- (7) Guy Fawkes was arrested in a cellar beneath the House of Lords with thirty-six **barrel of gunpowder**.

- (8) According to his manifesto, Flanagan was enraged at this woman. He was a **powder keg** just waiting to go off at her.

It appears that in some cases a learner's connection to their native language could hinder their ability to translate accurately. This is because they may be fixated on certain linguistic patterns that are specific to their native language (something like the functional fixation stated by Mandelblit (1995)). As a result, when translating into another language, they may choose words that are grammatically correct in that language but are not commonly used. They resort to what is possible at system-level (Halliday, 1991; Berber Sardinha, 2004) as opposed to norm-level choices (Coseriu, 1979).

Admissible solutions within the system also were a tendency verified by Mauranen (2008) when analyzing translator's production. This tendency was identified through a corpus-based methodology that examines the likelihood of linguistic patterns co-occurring.

Corpus analysis, as noted by Deignan (1997), is a valuable tool for exploring co-occurrences and can provide useful insights into how the metaphorical lexicon works, which can be beneficial for the translation practice.

Different Mapping Conditions (DMC) were employed in only 6% of the translation schemes, as shown in Table 2. This lower percentage in comparison to SMC (Same Mapping Conditions) strengthens the possibility of domain transposition from the source language (SL) to the target language (TL). It also indicates that, in relation to the analyzed domains, the languages would tend to have a common mapping. In every translation, the phraseological implementation is different (PhrDif), which is justified by the domain change. For instance, in (12), DMC is due to the connection of *trazer à tona* to the conceptual metaphor TRUTHS ARE OBJECTS SUBMERGED IN A CONTAINER, which is not observed in *bring to the table* (12):

- (9) *Aí está a outra verdade, está mais Heideggeriana, que Schwartz trouxe à tona: a maioria de nós perdeu a trilha do que é arte nos dias de hoje.*

That's another truth, this even closer to Heidegger, that Schwartz **brought to the table**: most of us has lost the track of what is art nowadays.

When learners attempt to use a domain distinct from the one in the source language, they may experience difficulties, as seen when translating *impermeável às críticas* to *unshakable to the criticism*, a non-conventional metaphorical phraseology (13):

- (10) But Temer stands on his feet and is **unshakable to the criticism** of those who call the legitimacy and harshness of his acts into question.

Regarding the relationship between PhrSim (41%) and PhrDif (51%), there appears to be no significant difference between them. However, there is more PhrDif because the DMC category is composed of a different lexical implementation. When these categories are analyzed based on the conventionality criteria, the percentage of PhrSim/MPhrConv is 25% and the percentage of PhrDif/MPhrConv is 25%. If we look only at PhrSim, we realize that when the learner chooses cognates or *prima facie* translations, they have a slightly higher chance of arriving at a conventional phraseology. The amount of PhrSim/MPhrLConv and MPhrNConv is 16%. And the percentage of PhrDif/MPhrLConv and PhrDif/MPhrNConv is 26%.

The division of domains was made to organize and group phraseologies for analysis. The purpose was not to evaluate how translations would take place in each domain or the challenges faced by learners in each one. However, the container, motion/direction, and control domains were the most productive in terms of metaphorical phraseologies. This may be because they have imagetic schemes that organize spatial experiences, making them more commonly found in languages (Lakoff, 1987; Kövecses, 2010). It is important to note that there was no specific criterion to delimit the number of domains, as they were organized based on phraseologies raised without a specific categorization in mind.

After analyzing Table 2, it was observed that the following domains were less difficult, as they had a higher percentage of MPhrConv in comparison to the other categories combined: container (56%), control (66%), war/physical conflict (60%), and games (70%). On the other hand, the most challenging domains were construction (59%), money/economic transaction (85%), play (61%), and other domains (61%). The plant domains accounted for 50% of MPhrConv and 50% of the other categories combined.

The research findings, with a predominance of SMC translations, are consistent with the results of similar studies. For example, Kövecses (2005) analyzed the translations of idioms from Hungarian to English and found that the most common strategy used was the maintenance of the conceptual metaphor. Martins (2008) points out a high level of conceptual similarity in metaphors used in abstracts written in Portuguese and translated into English. Additionally, Rodríguez Marquez (2010) found no cultural differences in terms of conceptual understanding when analyzing translation patterns in specialized language (economy) between American English and Mexican Spanish. Baiocco and Siqueira (2018) conducted a study on the use of conceptual mappings in literary translation from Portuguese to English. They found out that, despite some expressions being translated literally, there were still many similarities in the mappings used. This suggests that, from a conceptual perspective, the languages share many similarities in how they categorize reality. This observation supports the idea that there may be more conceptual approximations than differences between these languages.

Researchers who discuss the challenges of translating conceptual metaphors, such as Al-Haswani (2007), Kövecses (2014), Maalej (2008), and Mandelblit (1995), suggest that learners should be aware of the cultural and conceptual differences between languages to achieve a satisfactory translation. However, in addition to the difficulties posed by conceptual systems, there is also the issue of lexical choice that can influence the translation process.

Based on the data investigated, it appears that the main source of difficulty for language learners is not necessarily caused by the difference or lack of correlation between conceptual systems, as stated in Mandelblit (1995) and Kövecses (2014), or the lack of exposure to the conceptual system of the target language, as affirmed by Danesi (2004). Instead, this research suggests that the awareness of the conceptual mappings seems to be of secondary importance since only 6% of the phraseologies have DMC. The real problem lies on the lexical level or, more precisely, on the lexicon-semantic aspect, even in the face of common conceptual mappings.

The lexicon-semantic influence is apparent in phraseologies when assessing the conventionality parameter (MPhrConv/MPhrLConv and MPhrNConv). Regardless of the mapping condition or phraseological implementation, 50% of translations are MPhrConv, 22% are MPhrLConv, and 20% are MPhrNConv. At first glance, the MPhrConv numbers are high, indicating the learners' ability to suggest conventional choices. However, when considering less or non-conventional choices together, 42% of phraseologies signal a lack of familiarity among some learners with the metaphorical lexicon commonly used in the English language. Examples of non-conventional metaphorical phraseologies with SMC are presented from 14 to 17, while conventional ones, taken from the English reference corpus (*English Web 2015-enTenTen15*) are presented in parentheses.

- (11) *Nem sequer temos boas perspectivas para, num futuro próximo, dar o **salto de qualidade** de que o Brasil tanto necessita.*

They don't even have good perspectives to, in the near future, take a **leap of quality** Brazil needs so badly (**quantum leap** - 6,011 occurrences, *T-score* 77,50 *MI* 11,09 and 8,76).

- (12) *Embora alguns usuários possam imaginar que sejam capazes de controlar o consumo, cedo ou tarde, descobrem que, de fato, já não são **senhores de si** próprios.*

Although some users can think that they are capable of controlling the consume, sooner or later, they find out they are not indeed **their own owner (be your own master** – 2,111 occurrences, T-score 18,98, MI 2,96 and LogDice 5,09).

- (13) *Os anos de abuso descontrolado de recursos naturais [...] vinham matando rios e nascentes. A natureza cobrou seu preço.*

The years of uncontrolled abuse of natural resources [...] had been killing rivers and headwaters. Nature has **charged its price (takes its toll** – 31,835 occurrences, T-score 176,83 MI 6,81 and LogDice 5,51).

- (14) *Transcrevo o depoimento de um adicto recuperado. Ele fala com a força e a sinceridade de quem esteve no fundo do poço [...].*

I transcript the testimonial of a recover addicted. He speaks with strength and sincerity whom was in **bottom of the shaft (rock bottom** – 6,418 occurrences, T-score 79,38 MI 6,78 and LogDice 6,66).

Based on the analysis carried out, it appears that the awareness of the linguistic aspect has more influence than the awareness of the cognitive aspect. In line with Deignan, Gabrys e Solska's (1997) argumentation, having the same conceptual metaphor does not necessarily mean that the linguistic expression will be the same in different languages. As such, decoding of conceptual metaphor would not be a problem per se “however, the exact words and phrases which express this conceptual link in L2 cannot be guessed by reference to L1, so these need to be discussed and learned for encoding purposes” (Deignan, Gabrys; Solska, 1997, p. 353). Additionally, according to Kövecses (2014), the influence of linguistic and sociocultural knowledge in the formation of figurative expressions can lead to differences in the lexical level, even in the presence of similar cognitive processes.

The main conclusion of this research is that unawareness of the conventional metaphorical lexicon can turn the translation process into a more complex one. Theoretically, the linguistic structure leads to the conceptual metaphor – the learners' choices confirm this assumption –, however, even in the face of similar conceptual systems, it's important for learners to realize that not all lexical choices are widely accepted by a language community and may not be appropriate in every context.

This assertion supports Philip's (2010) viewpoint on the importance of phraseological and collocational patterns in expressing a conceptual arrangement. However, Philip argues that imprecise linguistic knowledge could result in unclear conceptual understanding. As the data showed us, linguistic imprecision can lead to an intelligible conceptual projection. The use of an unusual vocabulary could lead to collocation incongruences (Philip, 2010), which can cause semantic issues. Nevertheless, if a learner uses conventional metaphorical phraseology, they can automatically access the conceptual system of a language without any negative impact on their understanding. Therefore, although conceptual influence and metaphorical competence are important factors in learning a foreign/second language (Danesi, 1994, 2008; Littlemore; Low, 2006a), they are not enough conditions to produce conventional linguistic forms (Philip, 2010).

The findings of this research ratify Philip's observation that not knowing the association patterns of a word is the most responsible factor for incongruent constructions in other languages. They also ratify Boer's (2001) statement that understanding the conceptual struc-

ture doesn't necessarily affect conventional linguistic realization. In a parallel with the idiom principle (Sinclair, 1991), the more a learner is exposed to the conventional phraseology of a language, the more likely they are to internalize the phraseological structures of that language. As a result, they can access the conceptual system spontaneously, which enriches their linguistic and metaphoric competence, and improves their daily use of the language, whether in oral communication or translation.

The phraseological repertoire analyzed legitimates a learner profile that has an innocent view (Fillmore, 1979) of lexis conventional aspects (Tagnin, 2013, Orenha-Ottaiano, 2004; 2009; 2012). The research highlights a gap in the teaching of foreign/second language and translation, namely the need to focus on developing a curriculum that prioritizes the phraseometaphorical lexicon based on linguistic evidence. This is particularly relevant for Brazilian translator learners. Thus, the study emphasizes the importance of teaching conventionality in translation and calls for pedagogical solutions to improve the phraseometaphorical competence of learners.

5 Final remarks

This paper explored the findings of a study whose goal was to evaluate how learners react to the translation of metaphorical phraseologies. It focused on identifying similarities and differences in the mapping conditions between the Portuguese and English language, as well as examining the conventionality of the lexicon-semantic choices proposed, to see if these factors would cause difficulties in translation. The results showed that the conceptual aspect was less problematic, as the learners were able to identify the metaphors and keep them in the target language. Phraseologies with SMC (Similar Mapping Condition) accounted for 86% of the translations. On one hand, phraseologies with SMC and DMC (Different Mapping Conditions) guaranteed conventional linguistic realization in 50% of the translations. Conversely, the preservation of the metaphor was not always achieved at the norm level, as 42% of the translations were classified as less or non-conventional metaphorical phraseologies. In summary, the learners were able to retain the semantic-cognitive unit, but the main challenge was linguistic.

The lexicon-metaphorical composition of the translations is explained by the positive or negative influence of the mother tongue. This influence affects both the cognitive and linguistic-cultural aspects. The translator might use a conceptual and linguistic metaphor that exists in the target language, based on the structure of the source language. This is due to the cognitive-linguistic and cultural similarities between the two languages. Alternatively, the translator might keep the metaphor but use an uncommon wording. It can be concluded, therefore, confirming what was reported by Boers (2000) and Philip (2010), that understanding conceptual knowledge is not enough. It is equally important to understand how the lexicon is associated with conventional phraseological patterns while communicating in a foreign language. Otherwise, there is a risk of creating semantic problems due to lexical choices that are not commonly used in the target language.

As reported, the cognitive translation hypothesis establishes that differences in the mapping conditions can make translation difficult (Mandelblit, 1995). When evaluating metaphorical phrases, the research did not confirm the original hypothesis. Instead, it showed that

difficulties in translation, caused by the use of less common or unconventional lexicon, may not be due to different mapping conditions, as they take place even within a similar mapping condition. The research focused on the translation product, not on the time that learners spent on translation. Based on the cognitive translation hypothesis and subsequent works, an organogram with translation schemes was created. This organogram was used to examine the mapping conditions, linguistic implementation, and conventional characteristics of phraseologies, which helped identify translation tendencies.

The phraseometaphorical competence of translator learners can be improved by corpora as a tool for teaching metaphorical phraseologies. This can be achieved by using platforms like the one used in this research. The data-driven learning approach (Boulton, Tyne, 2013; Johns, 1991) can also be helpful and can promote learning autonomy in students, allowing them to explore corpora with guidance from their teacher. By analyzing concordance lines, students can identify word patterns and discover which words commonly go together in metaphorical contexts. They can compare their own intuitions with what is revealed by corpora. Authors such as Boers (2000), Charteris-Black (2002), Danesi (1994, 2008), Littlemore and Low (2006a), have made valuable contributions to the study of metaphor and can provide insights on how to implement a lexicon-metaphor based curriculum in the classroom. However, further research is needed to develop and deepen our understanding of this subject.

Supporting Agencies

I am grateful to the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) for funding this research through the Proex and Print programs. I am also thankful to the School of Education and Lifelong Learning at the University of East Anglia for waiving the school fees, which made it possible to develop this research during my PhD exchange program.

References

- AL-HASNAWI, A. R. A cognitive approach to translating metaphors. *Translation Journal*, v. 11, n. 3, 2007. Available from: <http://translationjournal.net/journal/41metaphor.htm>. Acess on: Mar. 2019.
- ALVES, F.; TAGNIN, S. E. O. corpora e ensino de tradução: o papel do automonitoramento e da conscientização cognitivo-discursiva no processo de aprendizagem de tradutores novatos. In: VIANA, V.; TAGNIN, S. (org.). *Corpora no ensino de línguas estrangeiras*. São Paulo: HUB Editorial, 2010. p. 189-203.
- ARDUINI, S. Metaphor, translation, cognition. In: MILLER, D.; MONTI, E. (ed.). *Tradurre figure/translate figurative language*. Emilia-Romagna: Università di Bologna, 2014. p. 41-52.
- BAIOCCO, L.; SIQUEIRA, M. Como se traduz metáfora? uma análise com base na teoria da metáfora conceitual. *Linguagem em foco*, Fortaleza, v. 10, n. 2, p. 79-90, 2018. Available from: <https://revistas.uece.br/index.php/linguagememfoco/article/view/1107>. Acess on: Oct. 15, 2019.
- BERBER SARDINHA, T. *Linguística de corpus*. Barueri: Manole, 2004.

- BERBER SARDINHA, T. Metáforas e linguística de corpus: metodologia de análise aplicada a um gênero de negócios. *Revista Delta*, São Paulo, v. 27, p. 1- 20, 2011a. DOI: <http://dx.doi.org/10.1590/S0102-44502011000100001>.
- BERBER SARDINHA, T. *Corpus linguistics in South America*. In: VIANA, V.; ZYNGIER, S.; BARNBROOK, G. (ed.). *Perspectives on corpus linguistics*. Amsterdam and Philadelphia: John Benjamins, 2011b. p. 29 - 45.
- BERBER SARDINHA, T. An assessment of metaphor retrieval methods. In: MACARTHUR, F. et al. (org.). *Metaphor in use: context, culture, and communication*. Amsterdam: John Benjamins, 2012. p. 21-50.
- BERBER SARDINHA, T. Metáforas da economia no dicionário de colocações do português brasileiro: uma análise multidimensional baseada em corpus. *Filologia e Linguística Portuguesa*, São Paulo, v. 18, n. 1, p. 175-198, 2016. DOI: <https://doi.org/10.11606/issn.2176-9419.v18i1p175-198>.
- BOERS, F. Metaphor awareness and vocabulary retention. *Applied Linguistics*, Oxford, v. 21, n. 4, p. 553-571, 2000. DOI: <https://doi.org/10.1093/applin/21.4.553>.
- BOULTON, A.; TYNE, H. Corpus linguistics and data-driven learning: a critical overview. *Bulletin Suisse de linguistique appliquée*, v. 97, p. 97-118, 2013. Available from: <https://hal.science/hal-01850687v1>. Access on: Oct. 10, 2019.
- BREZINA, V.; McENERY, T.; WATTAM, S. Collocations in context: a new perspective on collocation networks. *International Journal of Corpus Linguistics*, Amsterdam, v. 20, n. 2, p. 139-173, 2015. DOI: <https://doi.org/10.1075/ijcl.20.2.01bre>.
- CHARTERIS-BLACK, J. Second language figurative proficiency: a comparative study of Malay and English. *Applied Linguistics*, Oxford, v. 23, n. 1, p. 104 -133, 2002. DOI: <https://doi.org/10.1093/applin/23.1.104>.
- CORPAS PASTOR, G. *Manual de fraseología española*. Madrid: Gredos, 1996.
- COSERIU, E. *Teoria da linguagem e linguística geral*. Tradução de Agostinho Dias Carneiro. São Paulo: Edusp, 1979.
- DANESI, M. Research on metaphor and the teaching of Italian. *Italica*, Illinois, v. 71, n. 4, p. 453-464, 1994. DOI: <https://doi.org/10.2307/479665>.
- DANESI, M. Conceptual errors in second-language learning. In: KNOP, S. de.; RYCKER, T. de. (eds.). *Cognitive approaches to pedagogical grammar*. Berlin, New York: Mouton de Gruyter, 2008. p. 231- 256.
- DEIGNAN, A. *A corpus-based study of some linguistic features of metaphor*. 1997. 387 f. Tese (Doutorado) – Faculdade de Artes, Universidade de Birmingham, Birmingham, 1997.
- DEIGNAN, A., GABRÝS, D., SOLSKA, A. Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal*, Oxford, v. 5, n. 4, p. 352-360, 1997. DOI: <https://doi.org/10.1093/elt/51.4.352>.
- DEIGNAN, A. *Metaphor and corpus linguistics*. Amsterdam: John Benjamins, 2005.
- DURRANT, P.; SCHMITT, N. To what extent do native and non-native writers make use of collocations? *IRAL*, Berlin, v. 47, n. 2, p. 157-177, 2009. DOI: <https://doi.org/10.1515/iral.2009.007>.
- EVERT, S. *The statistics of word cooccurrences: word pairs and collocations*. 2005. 353 f. Tese (Doktors der Philosophie) – Institut für maschinelle Sprachverarbeitung, University of Stuttgart, 2005.

- FILLMORE, C. J. Innocence: a second idealization for linguistics. *Berkeley Linguistic Society*, v. 5, p. 63-76, 1979. DOI: <https://doi.org/10.3765/bls.v5i0.3255>.
- FRANKENBERG-GARCIA, A. Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes*, v. 35, p. 93-104, 2018. DOI: <https://doi.org/10.1016/j.jeap.2018.07.003>.
- GIBBS, R. W. *Embodiment and cognitive science*. Cambridge: Cambridge University Press, 2006.
- GRANGER, S.; PAQUOT, M. Disentangling the phraseological web. In: GRANGER, S.; PAQUOT, M.; MEUNIER, F. (ed.). *Phraseology: an interdisciplinary perspective*. Amsterdam: John Benjamins, 2008. p. 27 – 50.
- GRANGER, S.; LEFER, M.-A. MUST: A collaborative corpus collection initiative for translation teaching and research. In: GRANGER, S., LEFER, M.-A.; PENHA-MARION, L. (eds), *Book of Abstracts. Using Corpora in Contrastive and Translation Studies Conference (5th edition)*. CECL Papers 1. Louvain-la-Neuve: Centre for English Corpus Linguistics/Université catholique de Louvain, 72-73, 2018. Available from: <https://uclouvain.be/en/research-institutes/ilc/cecl/cecl-papers.html>. Access on: Mar. 10, 2019.
- GRUPO PRAGGLEJAZ. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and symbol*, Philadelphia, v. 22, n. 1, p. 1–39, 2007. DOI: <https://doi.org/10.1080/10926480709336752>.
- HALLIDAY, M. A. K. Corpus studies and probabilistic grammar. In: AIJMER, K.; ALTENBERG, B. (ed.). *English corpus linguistics: studies in honour of Jan Svartvik*. London: Longman, 1991.
- IRUJO, S. Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *Tesol Quarterly*, v. 20, n. 2, p. 287–304, 1986. DOI: [10.2307/3586545](https://doi.org/10.2307/3586545).
- JOHNS, T. Should you be persuaded: two examples of data-driven learning. In: JOHNS, T.; KING, P. (ed.). *Classroom Concordancing*. Birmingham: English Language Research Journal, 1991. p. 1-16.
- KILGARRIFF et al. The sketch engine: ten years on. *Lexicography: Journal of ASIALEX*, Sheffield, v. 1, n.1, p. 7-36, 2014. DOI: <https://doi.org/10.1007/s40607-014-0009-9>.
- KÖVECSES, Z. *Metaphor in culture: universality and variation*. Cambridge: Cambridge University Press, 2005.
- KÖVECSES, Z. *Metaphor: a practical introduction*. New York: Oxford University Press, 2010.
- KÖVECSES, Z. Conceptual metaphor theory and the nature of difficulties in metaphor translation. In: MILLER, D.; MONTI, E.; (ed). *Tradurre figure/translating figurative language*. Quaderni del CeSLiC, Atti di Convegni, 2014. p. 25 – 39.
- LAKOFF, G. *Women, fire and dangerous things*. Chicago: University of Chicago Press, 1987.
- LAKOFF, G.; JOHNSON, M. *Metaphors we live by*. Chicago: The University Press of Chicago, 2003 [1980].
- LAKOFF, G.; ESPENSON, J.; GOLDBERG, A. *Master metaphor list (draft copy)*. Berkeley: University of California, 1989.
- LITTEMORE, J.; LOW, G. Metaphoric Competence, second language learning, and communicative language ability. *Applied Linguistics*, Oxford, 2006a, v. 27, n. 2, p. 268 -294. DOI: <https://doi.org/10.1093/applin/aml004>.
- LITTEMORE, J.; LOW, G. *Figurative thinking and foreign language learning*. New York: Palgrave Macmillan, 2006b.

- LITTLEMORE, J et al. An investigation into metaphor use at different levels of second language writing. *Applied Linguistics*, Oxford, v. 35. n.2, 2014, p.117-144. DOI: <https://doi.org/10.1093/applin/amtoo4>.
- MAALEJ, Z. Translating metaphor between unrelated cultures: a cognitive pragmatic perspective. *Translation Journal*, Sayyab, v. 1, p. 60 -61, 2008.
- MANDELBLIT, N. The cognitive view of metaphor and its implications for translation theory. *Translation and Meaning*, v. 3, p. 483-495, 1995.
- MARTINS, L. de M. *Identificação e tradução de metáforas linguísticas em abstracts da esfera acadêmica: uma análise baseada em Linguística de Corpus*. 2008. 199 f. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) — Pontifícia Universidade Católica, São Paulo, 2008.
- MAURANEN, A. Universal tendencies in translation. In: ANDERMAN, G.; ROGERS, M. (eds). *Incorporating corpora: the linguist and the translator*. Clevendon: Multilingual Matters, 2008, p. 32-48.
- MCENERY, T. *Swearing in English: bad language, purity and power from 1586 to the present*. Abington: Routledge, 2006.
- NESSELHAUF, N. *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company, 2005.
- ORENHA-OTTAIANO, A. *A compilação de um glossário bilíngue de colocações, na área de jornalismo de negócios, baseado em corpus comparável*. 2004. 246 f. Dissertação (Mestrado em Estudos Linguísticos e Literários) — Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2004.
- ORENHA-OTTAIANO, A. *Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não juramentado*. 2009. 282 f. Tese (Doutorado em Estudos Linguísticos) — Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2009.
- ORENHA-OTTAIANO, A. English collocations extracted from a corpus of university learners and its contribution to a language teaching pedagogy. *Acta Scientiarum: Language and Culture*, Maringá, v. 34, n. 2, p. 241-251, 2012a. DOI: <https://doi.org/10.4025/actascilangcult.v34i2.17130>.
- PHILIP, G. Drugs, traffic, and many other dirty interests: metaphor and the language learner. In: LOW, G.; TODD, Z.; DEIGNAN, A.; CAMERON, L. (ed.). *Researching and applying metaphor in the real world*. Amsterdam: John Benjamins, 2010. p. 63-80.
- PHILIP, G. *Colouring meaning: collocation and connotation in figurative language*. Amsterdam: John Benjamins Publishing, 2011.
- ROCHA, J. M. P. *Fraseologia jurídico-comercial e proposta de um glossário de colocações especializadas trilíngue baseado em corpus*. 2017. 292 f. Dissertação (Mestrado em Estudos Linguísticos) — Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2017.
- ROCHA, J. M. P. *Tradução de fraseologismos metafóricos do português para o inglês: um estudo de corpus de aprendizes brasileiros*. 2020. 257f. Tese (Doutorado em Estudos Linguísticos — Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2020a.
- ROCHA, J. M. P.; VIANA, V.; ORENHA-OTTAIANO, A. Tradução de fraseologismos metafóricos: contribuições teórico-metodológicas da linguística de corpus. *Revista da ABRALIN*, [S. l.], v. 19, n. 1, p. 1-26, 2020b. DOI: [10.25189/rabralin.v19i1.1697](https://doi.org/10.25189/rabralin.v19i1.1697)

- RODRÍGUEZ MARQUEZ, M. de M. *Patterns of translation of metaphor in annual reports in American English and Mexican Spanish*. 2010. 146 f. Tese (Doctor of Philosophy) – Faculty of Arts and Human Sciences, University of Surrey, 2010.
- RUIZ GURILLO, E. *Aspectos de fraseología teórica española*. Valencia: Universitat de València, 1997.
- SAUSSURE, F. *Curso de linguística geral*. São Paulo: Cultrix, 2012.
- SEMINO, E. *Metaphor in discourse*. Cambridge: Cambridge University Press, 2008.
- SINCLAIR, J. *Corpus, concordance, collocation*. Oxford: Oxford University Press, 1991.
- SINCLAIR, J. The search for units of meaning. In: *CICLE de Conferències 95-96: Lèxix, corpus i diccionaris*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 1996. p. 97-107.
- SINCLAIR, J.; JONES, S.; DALEY, R. English lexical studies: report to OSTI on project C/LP/08: new edition. In: KRISHNAMURTHY, R. (ed.). *English collocation studies: the OSTI report*. Birmingham: Birmingham University Press, 2003. p. 2-138.
- STEEN, G. *Finding metaphor in grammar and usage: a methodological analysis of theory and research*. Amsterdam: John Benjamin Publishing, 2007.
- STEEN, G. et al. Metaphor in usage. *Cognitive Linguistics*, v.21 n. 4, p. 765 – 796, 2010. DOI: <https://doi.org/10.1515/cogl.2010.024>.
- STUBBS, M. *Words and phrases: corpus-based studies of lexical semantics*. Oxford: Routledge, 2001.
- TAGNIN, S. E. O. *O jeito que a gente diz: expressões convencionais e idiomáticas*. São Paulo: Disal, 2013.
- TAHERI-ARADALI, M.; BAGHERI, M.; EIDY, R. Towards a new model of metaphor translation: a cognitive approach. *Iranian Journal of Translation studies*, Mashhad, v. 11, n. 41, p. 35-52, 2013.
- TOGNINI-BONELLI, E. *Corpus linguistic at work*. Amsterdam: John Benjamins, 2001.
- VEREZA, S. “Metáfora é que nem...”: cognição e discurso na metáfora situada. *Signo*, v. 38, n. 65, p. 2-21, jul./dez., 2013. DOI: <http://dx.doi.org/10.17058/signo.v38i65.4543>.
- VEREZA, S. Cognição e sociedade: um olhar sob a óptica da linguística cognitiva. *Linguagem em(Dis)curso – LemD*, v. 16, n. 3, p. 561-573, set./dez. 2016. DOI: <https://doi.org/10.1590/1982-4017-160303-0416d15>.
- ZULUAGA, A. *Introducción al estudio de las expresiones fijas*. Frankfurt: Peter D. Lang, 1980.

A variação da ordem pronominal em complexos verbais na variedade moçambicana do Português

The Variation of Pronominal Order in Verb Complexes in the Mozambican Variety of Portuguese

Bento Orlando Mutoba

Universidade Estadual de Feira de Santana
(UEFS) | Feira de Santana | BA | BR
CAPES

bentomutoba99@gmail.com
<https://orcid.org/0009-0007-9629-0282>

Norma Lucia Fernandes de Almeida

Universidade Estadual de Feira de Santana
(UEFS) | Feira de Santana | BA | BR
normauefs@gmail.com
<https://orcid.org/0000-0002-3369-4251>

Resumo: Este artigo analisa a variação da ordem pronominal em complexos verbais na variedade moçambicana, com intuito de compreender o seu nível de ocorrência e encaixe na matriz linguística e social. Constitui uma das preocupações da sociolinguística - na sua interseção com diversas perspectivas de abordagem dos estudos de linguagem - o entendimento do encaixamento do fenômeno em variação a fim de se descrever as tendências que dele se observa. Para tanto, este estudo ancora-se no aporte da Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 1968 [2006]), tendo o seu *corpus* constituído com base em falas extraídas de 3 vídeos de 3 programas televisivos da cidade de Maputo (Sul) e cidade de Nampula (Norte). Considerou-se um total de 8 informantes (4 por cada cidade) para constituição da amostra da pesquisa. Os resultados apontam que a variação da ordem pronominal em complexos verbais no Português falado em Moçambique, que se reflete por meio da produtividade da variante intra-CV com próclise ao verbo temático infinitivo independentemente do contexto morfossintático, apresenta um grau de encaixamento bastante significativo na matriz linguística e social, o que nos leva a pensar numa possível mudança em curso na variedade moçambicana em detrimento da norma europeia que vigora como padrão em Moçambique.

Palavras-chave: variedade moçambicana; colocação pronominal; complexos verbais; encaixamento.



Abstract: This paper analyses the variation of the pronominal order in verbal complexes in the Mozambican variety of portuguese, with the aim of understanding its occurrences and how it fits into the linguistic and social matrix. One of the concerns of sociolinguistics - in its intercession with various perspectives of language studies - is to understand the embedding of the phenomenon in variation in order to describe the tendencies that can be observed in it. To this end, this study is based on the Theory of Variation and Change (Weinreich; Labov; Herzog, 1968 [2006]), and its corpus is made up of speeches extracted from 3 videos of 3 television programs from Maputo city (South) and Nampula City (North). A total of 8 informants (4 for each city) were considered to constitute the research sample. The results show that the variation in pronominal order in verb complexes in Mozambican Portuguese, reflected in the productivity of the intra-CV variant with proclisis to the infinitive thematic verb regardless of the morphosyntactic context, shows a very significant degree of embedding in the linguistic and social matrix, which leads us to think of a possible change in progress in this variety when compared to the European one, which is taken as the standard norm in Mozambique.

Keywords: Mozambican portuguese; pronominal collocation; verb complexes; embedding.

1 Introdução

Os estudos sociolinguísticos são desenvolvidos atendendo a vários pressupostos que permitem sustentar, sobre o “objeto língua”, o ponto de vista de que as dinâmicas de qualquer língua natural podem ser mais bem compreendidas considerando a sua relação com a sociedade, em que estão envolvidos vários aspectos que ajudam a explicar a diversidade da língua falada/sinalizada, observada, descrita e analisada em seu contexto social, isto é, em situações reais de uso.

A abordagem dos estudos de linguagem exige atenção a vários aspectos, como o problema de “encaixamento”, por exemplo, fundamental na sociolinguística. Esse conceito foi abordado por Weinreich, Labov e Herzog (1968), para explicarem que qualquer fenômeno sob investigação deve ser visto como encaixado no sistema linguístico, elucidando a natureza e extensão desse encaixamento na estrutura linguística, como também na estrutura social referente ao encaixamento social e geográfico.

Os autores anteriormente citados acreditam que, no encaixamento da variação, o contexto linguístico que favorece um determinado tipo de mudança desencadeia outros, em possíveis relações em cadeia, por isso uma análise restrita é insuficiente para dar conta da mudança.

Desta feita, o encaixamento deve ser observado tanto na *estrutura linguística* - que pode se dar da covariação com elementos linguísticos e extralinguísticos, como na *estrutura social*, havendo a necessidade de se olhar para uma determinada comunidade de fala para entender como o ambiente social influencia na variação, o que pode levar à mudança linguística.

Posto isto, entendemos que a variedade moçambicana do Português, embora bastante recente devido ao pouco tempo que Moçambique conseguiu a sua independência do jugo colonial, vai cada vez mais ganhando uma dinâmica própria. Apesar de Moçambique ser um país com cultura de língua padrão no qual a preocupação em se ajustar aos modelos da norma-padrão é grande, tem sido comum, na modalidade oral dos falantes, registrar sentenças como:

- (1) *Sabe, desse jeito, a ferida **vai se apresentar** com dificuldade para sarar devido à fragilidade das células que não conduzem o açúcar para aquela região da ferida. E numa fase avançada o paciente pode desenvolver problemas graves de vista devido a dilatação da pupila.* (Fi, I-FN)
- (2) *A minha não é sentença, a minha revolta **não pode me dar** cadeia, a minha revolta é de um filho que não tem pão.* (Fii, I-CM)
- (3) *[...] se **tiverem que me prenderem** que me prendam aqui, as cadeias não foram feitas para animais.* (Fii, I-CM)

Nas sentenças (1) a (3) da variedade do PM há, nos complexos verbais, colocação de clíticos pronominais em contextos não licenciados pela norma europeia: (1) ocorrência do clítico pronominal na posição intracomplexo verbal (intra-CV) com próclise ao verbo principal (v2) em contextos em que o complexo verbal (adiante CV) é precedido por apenas sintagma nominal (SN) com função gramatical de sujeito.

Em (2) temos na 2^a oração coordenada assindética uma categoria negativa (não) na posição pré-CV e clítico no interior do complexo verbal (CV), com próclise ao verbo infinitivo temático, contexto em que a norma da variedade europeia não permitiria a ocorrência do clítico no interior do CV devido à presença de um elemento proclisador (não), e havendo, assim, apenas a possibilidade de próclise (cl v1 v2) ao verbo auxiliar ou ênclide ao verbo principal (v1 v2-cl).

Na sentença (3) temos o clítico na mesma posição que se encontra nas sentenças (1) e (2), todavia diferindo-se das demais pelo fato de ser um contexto frásico reconhecido pela norma padrão europeia que vigora em Moçambique.

Assim, na variedade moçambicana parece haver uma crescente tendência, em complexos verbais, de se favorecer uma ordem (v1 cl v2), independentemente do contexto, diferente daquela que é produtiva na variedade europeia (v1-cl v2) de acordo com Vieira (2016).

Nesse prisma, buscamos, no geral, analisar a variação da ordem pronominal em complexos verbais no Português falado em Moçambique. De modo específico, buscamos descrever a colocação pronominal em complexos verbais no Português falado em Moçambique; avaliar o nível de ocorrência e de encaixe da variação da ordem pronominal em complexos verbais na matriz linguística e social no Português de Moçambique (adiante PM).

Uma das nossas motivações alinha-se ao fato de a ordem pronominal em estruturas verbais complexas ser pouco descrita nos estudos sobre o tema – que se ocupam, em sua

maioria, das lexias verbais simples (Vieira, 2023). Metodologicamente, importa referir que se trata de um estudo descritivo e exploratório e, para tanto, apropriamo-nos do *corpus* constituído com base em falas extraídas de 3 vídeos de 3 programas televisivos da cidade de Maputo (Sul) e cidade de Nampula (Norte); onde, em função da possibilidade do controle do perfil, consideramos um total de 8 informantes (4 por cada cidade) para a amostra da pesquisa.

Em estruturas complexas é desafiador identificar a posição exata do clítico quando ocorre no interior do CV, à medida que ele pode estar hospedado na posição enclítica do v1 ou proclítica do v2. Em torno disso, observamos alguns pontos que fundamentam a nossa percepção sobre a ordem do clítico no interior do CV, como parâmetros acústicos (duração, intensidade), o modo da apresentação dos constituintes morfossintáticos, a posição dos elementos intervenientes no interior do CV.

2 Nota situacional sobre o Português em Moçambique

Moçambique – país localizado no sudeste do continente africano, África austral, e banhado pelo oceano Índico a leste e que faz fronteira com a Tanzânia ao norte; Malawi e Zâmbia a noroeste; Zimbábue a oeste e Essuatíni e África do Sul a sudoeste – é hoje um dos países africanos de expressão portuguesa tal como Cabo Verde, Angola, Guiné-Bissau, São Tomé e Príncipe, devido à presença dos portugueses nos seus territórios, que resultou na colonização que se refletiu tanto a nível sociocultural, ideológico, assim como linguístico; este último, por exemplo, se deu por meio da evangelização e por imposição do uso da língua portuguesa em diferentes contextos sociais, em detrimento das várias línguas que coabitam com a língua portuguesa nos países africanos.

De acordo com Ngunga (2012), estima-se a existência de mais de 20 línguas do extrato bantu faladas em Moçambique. Por conta dessa diversidade linguística, após o alcance da independência nacional em 1975, o português passou a ter o estatuto de língua oficial, de ensino e de unidade nacional.

Antes da independência, o português da metrópole era prescrito como norma padrão para toda a vida pública no seio da política de assimilação portuguesa. Nesse sentido, a política ultramarina na altura tencionava como objetivo a ‘longo prazo’ transformar todos os moçambicanos em cidadãos linguística e culturalmente semelhantes aos portugueses, e todo desvio era tido como erro e violação do “bom Português”. Após a independência, embora alguns linguistas como Ngunga (2004), Gonçalves (2005, 2010) e Mendes (2010) começem a considerar algumas particularidades do PM, principalmente a nível lexical, a variedade do português europeu continua como norma padrão em Moçambique; a recente saída dos portugueses do território nacional é apontada como uma das razões de ainda não haver condições para a oficialização de uma norma nacional.

Nesse sentido, Moçambique adota ainda a cultura de língua padrão em efeito da forte pressão da norma da variedade europeia do Português como modelo de língua portuguesa (LP). Essa cultura se dá até hoje pela crença na correção e na ideologia do padrão, sendo, normalmente, qualquer variedade alheia a norma padrão vista como errada, pois na consciência de muitos falantes só uma variante de uso da língua pode estar certa e acreditam que a língua existe apenas de forma padronizada (Mutoba e Almeida, 2023).

No entanto, mesmo vigorando a norma da variedade do PE como padrão em Moçambique, observamos com naturalidade a ocorrência de alguns desvios em quase todos falantes moçambicanos e contextos de uso da língua portuguesa (Gonçalves, 2010; Mutoba e Almeida, 2023; Nhatuve, 2017, Timbane (2012), Timbane 2017), tal é o caso do fenômeno da colocação dos clíticos pronominais em complexos verbais que parece não depender muito do contexto morfossintático ou frásico, diferentemente da variedade europeia que, por exemplo, é dependente de certos contextos, como a presença de elementos proclíticos na sentença. A seguir passamos a tratar dos padrões sobre a ordem dos clíticos pronominais em grupos verbais à luz da norma europeia.

3 A gramática dos clíticos pronominais em complexos verbais

No âmbito da colocação pronominal, importa referir que existem duas gramáticas autônomas, que podem ser tomadas como referência para a discussão sobre o tema: a da variedade europeia e a da variedade brasileira. Destarte, apresentamos os padrões sobre a ordem dos clíticos pronominais em grupos verbais à luz da norma europeia, tecendo algumas considerações da variedade brasileira.

Os pronomes clíticos, também designados de pronomes átonos ou clíticos especiais, “correspondem às formas átonas do pronome pessoal que ocorrem associados à posição dos complementos dos verbos” (Brito; Duarte; Mattos, 2003, p. 826-827). Eles caracterizam-se, de acordo com Mateus et al. (2003, p. 831), por não ter uma posição fixa relativamente ao seu hóspedeiro, podendo precedê-lo (próclise), seguí-lo (ênclide) ou ocorrer no seu interior (mesóclise).

Segundo Brito; Duarte e Mattos (2023), os pronomes clíticos têm um comportamento uniforme quanto aos padrões de colocação, isto é, todos eles exigem um hóspedeiro verbal, “o que se traduz numa vizinhança entre o clítico e uma forma verbal, finita ou não finita” (Brito; Duarte e Mattos, 2003, p. 847). E podem ocorrer à direita ou à esquerda do hóspedeiro, porém essas posições não se encontram em variação livre, significa que a sua variação depende de determinados contextos morfossintáticos. Desse modo, a seguir buscamos compreender como ocorre a hospedagem dos clíticos nos diferentes contextos possíveis na variedade do português europeu.

3.1 Variedade do Português Europeu

Na variedade do português europeu, de acordo com Brito; Duarte e Mattos (2003, p. 849-850), a posição enclítica é o padrão básico e marcado; e a posição proclítica é induzida por fatores de natureza morfossintáticas. Igualmente, autores como Carneiro (2016), Vieira, F. (2011), Vieira, F. (2016) e Vieira, S. e Vieira, F. (2018), ao explicarem a abordagem tradicional sobre a colocação dos clíticos pronominais, salientam que a ênclide ao verbo auxiliar ocorrerá quando não houver um contexto favorecedor da variante proclítica ao verbo auxiliar.

Nesse sentido, nas estruturas verbais simples, a ênclide ao verbo finito é regra geral, exceto em contextos que existam algumas condicionalidades que exijam a próclise. E essa propriedade se dá também com os verbos auxiliares nos grupos verbais quando há subida de

clítico¹. Sendo assim, é importante de alguma forma ter-se o conhecimento dos padrões de colocação dos clíticos em uso em construções de uma só forma verbal no PE.

Conforme já avançado *a priori*, nas descrições gramaticais do português europeu, a ênclise em estruturas verbais simples sempre é possível em contextos sem nenhuma categoria proclisadora, ainda que haja antecedência de algumas categorias sintagmáticas, como SN, seja em oração principal (*Sim, pedi-lhe que ele fizesse*) ou coordenada (*O crime é inextirpável, mas o combate dissuade-o e pune-o*).

Por seu turno, Martins (2016), ao estudar sobre a colocação dos pronomes clíticos em sincronia e diacronia, observa que o sistema do português europeu contemporâneo apresenta próclise e ênclise quer em frases finitas, quer em frases não finitas, com as duas ordens em distribuição complementar nas frases finitas, mas em variação livre em alguns tipos de orações infinitivas.

Assim, a próclise é normalmente possível em determinados contextos morfossintáticos ou frásicos, como em contextos nos quais há ocorrência de algumas categorias funcionando como atratores, tal é o caso de conjunções em subordinadas finitas, pois Martins (2016) explica que, se o clítico ocorrer dentro de uma oração subordinada fica na posição proclítica, todavia, se pertencer a uma oração principal, ocorre na posição enclítica.

Além disso, são também operadores da próclise as categorias negativas (*não, nunca, jamais*); quantificadores (*ambos, bastante, demasiado, demasiados, mais, menos, muito, pouco/s, raramente, suficientes, tal, tais, tão, todo/s, todo, tudo*); sintagmas QU- (*que, o que, quem, onde, quanto, como, quando, por que*); advérbios (*até, só, lá enfático, também, sempre, já, talvez*); marcadores de foco (inclusivos: *também, até, mesmo*; exclusivos: *apenas, só, somente, logo, antes; aspectuais: ainda, já, quase, mal, talvez*); marcadores de ênfase (*até, sempre, já*); dêiticos locativos (*aí, ali, aqui, cá, lá*). (Martins, 2016).

Existem também contextos que tornam possível a variação entre próclise e ênclise com o infinitivo simples, devido à natureza de algumas preposições que não são categóricas em relação à necessidade da ênclise ou próclise (de, a, com), significa que estas aceitam tanto a ênclise quanto a próclise (*Não parou de se queixar o tempo todo; Não parou de queixar-se o tempo todo*).

Relativamente às estruturas construídas por duas ou mais formas verbais com algum grau de integração sintático-semântica, e que constituem os chamados complexos verbais, buscamos referência nos estudos de Vieira, S. (2002); Vieira e Vieira, (2018) que estudaram a ordem dos clíticos pronominais na variedade oral do PE, tendo observado que a variante mais utilizada ou efetiva é a intra-CV com ênclise ao verbo auxiliar/v1 (*deixou-me a pensar no quarto; pode-me entregar essa lista?*). E além da variante enclítica a v1, também se registra com produtividade a ênclise à segunda forma verbal (*pode entregar-me essa lista?*), salvo os casos da existência de proclisadores, conforme Vieira e Vieira (2018) explicam:

A presença de elementos proclisadores, embora favoreça a próclise ao complexo (*que me pode dizer*), não apresenta atuação categórica, tendo havido também ênclise à primeira forma verbal (em complexos com gerúndio e particípio, como em *que vem-me visitando* ou *que tinha-me encontrado*) ou à segunda (em complexos com infinitivo, como em *que pode dizer-me*). Não havendo elemento proclisador, a ênclise – a v1 ou a v2-constitui opção natural. Clíticos imediatamente antes de v2 ficam restritos a complexos com *ter que/de* – o que não permite afirmar uma efetiva próclise a v2 (Vieira; Vieira, 2018, p. 284).

¹ Consiste na seleção de um verbo do qual o pronome clítico não é dependente para hospedeiro verbal (Brito; Duarte, 2003, p. 857), ou seja, quando o clítico se hospeda no verbo auxiliar, tanto posição enclítica, quanto na proclítica.

Entendemos com isso que, assim como nas construções com uma só forma verbal, a ênclise é produtiva na variedade do PE em contextos sem elementos proclisadores, sendo que nos grupos verbais a ênclise pode se dar, tanto no v1, quanto no v2, embora, diferentemente das dos padrões descritos pelas gramáticas, haja também algumas variações em contextos que se prevê uma determinada ordem.

Todavia, Vieira, F. (2011) esclarece que a variação da colocação dos clíticos pronominais em complexos verbais decorre também em função da forma em que o verbo principal aparece, pois em algumas formas nem todas as posições são possíveis, isto é, em orações com participípio e gerúndio, a posição pós-CV pode não ser admitida; normalmente, admite-se, na ausência de um atrator, a ênclise ao v1 e, na presença deste, admite-se a próclise. Assim, a autora descreve, ao interpretar a gramática de Mateus *et al.*, (2003), que:

Quando o verbo principal está no participípio ou no gerúndio, o clítico aparece obrigatoriamente proclítico (caso haja elemento “atrator” do pronome) ou enclítico ao verbo auxiliar (*Ele tinha-a chamado para a festa*). Nas construções com verbos semiauxiliares do tipo aspectual que selecionam infinitivos preposicionados, na presença de um elemento proclisador, há o favorecimento da subida do clítico com as preposições *a* e *de* (*Ele não lhe começou a dizer a verdade*), mas com a preposição *por*, isso não ocorre (*ele acabou por se esquecer do casamento*) (Mateus *et al.*, 2003, p. 25)

Nessa linha, é verdade afirmar que os complexos verbais com verbo principal no infinitivo é que admitem as quatro variantes possíveis (cl v1 v2, v1-cl v2, v1 cl v2, v1 v2-cl). É, justamente, nestas construções que a ênclise é possível no v1 ou no v2; e a próclise é possível ao verbo auxiliar quando houver antecedência de um atrator, e ao verbo principal infinitivo quando também estiver antecedido por um elemento proclisador, como é o caso das construções com as preposições *de* e *por*, nas quais ocorre obrigatoriamente a posição proclítica ao verbo principal (*Ele deixou **de lhe dizer** a verdade*).

Ainda na senda dos estudos da colocação de clíticos pronominais em grupos verbais, Vieira, F. (2011) na sua dissertação intitulada “A cliticização pronominal em lexias verbais simples e em complexos verbais no Português europeu oral contemporâneo: uma investigação sociolinguística”, explica que na modalidade oral do PE, em complexos verbais, a variante mais utilizada é a intra-CV, em que o clítico está, na maioria dos casos, enclítico a v1, sendo possível a próclise com a atuação dos atratores.

Por seu turno, Vieira, F. (2016) observa que na oralidade, a variante intra-CV é produtiva tanto na variedade europeia, como na variedade moçambicana e brasileira do Português. Porém, com ênclise a v1 na variedade europeia e moçambicana (embora nesta última a autora reconheça margens de dúvida) e, conforme se lê também em Martins (2016); Carneiro (2016); Vieira, F. (2016); Vieira e Vieira (2018); Araújo e Silva (2019), com próclise generalizada ao V2 na variedade brasileira, quer com as formas finitas quer com as formas não finitas do verbo, incluindo o infinitivo, o gerúndio e o participípio passado. Exceto em alguns casos cristalizados nas três variedades como: *isto hoje fez-me pensar... fez/fez-me pensar... muito; aí resta-nos saber onde é que está? trata-se* de olhar para frente e seguir.

Já Martins (2016) entende que as variedades orais africanas do português, incluindo a moçambicana, seguem em alguns casos a norma europeia, mas com “uma pequena margem de variação próclise/ênclise em contextos que no padrão europeu apenas permitem

uma das ordens" (Martins, 2016, p. 410). E nessas variedades é sonante ainda o problema levantado por Vieira, F. (2016) referente à dificuldade de determinar com firmeza, pela simples audição dos enunciados, que, em um enunciado como *<pode me dar>*, o *<me>* esteja ligado ao *<pode>* ou ao *<dar>*.

Assim, será importante entender o tipo dos verbos, clíticos entre outras variáveis envolvidas nos dados dessa pesquisa, para compreendermos o nível de ocorrência e encaixe da ordem pronominal em CV na matriz linguística e socialda na variedade moçambicana do português, através da metodologia que a seguir apresentamos.

4 Caminhos metodológicos

Este artigo baseia-se na Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 1968 [2006]), do campo da sociolinguística, que procura analisar e sistematizar as variações linguísticas usadas por uma comunidade de fala, tendo sempre em mobilização os fatores internos e externos que podem influenciar a dinâmica de qualquer língua.

Desse modo, Trindade (2021), ao interpretar a teoria de William Labov da variação e mudança, explica que:

Os sociolinguistas estudam a relação entre língua e sociedade, mostrando a variação e mudança linguística a partir dos pontos de vista diacrônico e sincrônico, entendendo que a língua possui um funcionamento dinâmico e não mecânico, articulando o comportamento linguístico e o social. Do ponto de vista diacrônico, o pesquisador estabelece dois momentos sucessivos de uma determinada língua, descrevendo-os e distinguindo as variantes que estão em desuso. Do ponto de vista sincrônico, aborda tomando por base três pontos de vistas: diatópico, diastrático e estilístico (Trindade, 2021, p. 52)

Assim, a sociolinguística busca analisar e sistematizar as variações linguísticas usadas por uma mesma comunidade de fala, tendo sempre em mobilização os fatores internos e externos que condicionam o funcionamento de qualquer língua natural, caracterizada pela dinâmica, tanto a nível sincrônico, assim como diacrônico.

Por sua vez, este estudo toma uma perspectiva sincrônica na medida que pretende descrever o fenômeno da colocação dos clíticos pronominais em complexos verbais a partir do presente referente a um contexto temporal específico.

O *corpus* do artigo foi constituído com base em falas extraídas de 3 vídeos de 3 programas televisivos, sendo um da cidade de Maputo (Sul) com duração de 2h e outros dois da cidade de Nampula (Norte) com a duração de 1h30min. No vídeo com duração de 2h considerou-se a fala de 4 participantes, e nos vídeos com duração de 1h30min, em um considerou-se a fala de três (3) participantes, e no outro considerou-se a fala de um participante em função da possibilidade do controle do perfil do informante. Desta feita, em função da possibilidade do controle do perfil, considerou-se um total de 8 informantes (4 por cada cidade) para a constituição da amostra da pesquisa.

A teoria laboviana considera que qualquer fenômeno sob investigação deve ser visto como encaixado no sistema linguístico, elucidando a natureza e extensão desse encaixamento na estrutura linguística; como também na estrutura social.

Nisso, é importante a mobilização de vários fatores internos e externos nos estudos de linguagem, para analisar e sistematizar as variações linguísticas usadas por uma comunidade de fala. Dessa feita, consideramos como fatores externos a região (cidade), sexo, idade e escolaridade para compreender a natureza de encaixamento do fenômeno dentro dessas comunidades de fala.

Tabela 1 – Distribuição das variáveis extralingüísticas/sociais.

	Cidade		Sexo		Idade		Escolaridade	
	Sul	Norte	M	F	Faixa-I	Faixa II	Média	Superior
Valores Absolutos	4	4	5	3	24-30	37-41	3	5
Percentuais	50%	50%	62,5%	37,5%	62,5%	37,5%	37,5%	62,5%

Fonte: Elaboração própria.

E controlamos as seguintes variáveis internas/linguísticas, inspirados pelos trabalhos de Vieira, S. (2002); Vieira, F. (2011); Vieira, F. (2016); i) o tipo de clítico, ii) tipo de oração, iii) presença ou não do elemento proclisador; iv) distância entre V-CL ou CL-V e um possível elemento antecedente, pessoa, tempo e; v) o tipo de estruturas verbais complexas.

A análise e interpretação dos dados fizemos fundamentalmente a partir dos valores absolutos e percentuais. É, ainda, pontual salientar que se trata de uma pesquisa exploratória, descritiva, quantitativa e qualitativa.

Ao estudar o fenômeno da ordem pronominal em estruturas complexas, na modalidade oral, surgem grandes debates sobre a posição exata do clítico no interior do CV. Segundo Vieira, F (2016) não se pode determinar, pela simples audição dos enunciados, que, em um enunciado como “*pode me dar*”, a forma “*me*” esteja ligada ao *<pode>* ou ao *<dar>*.

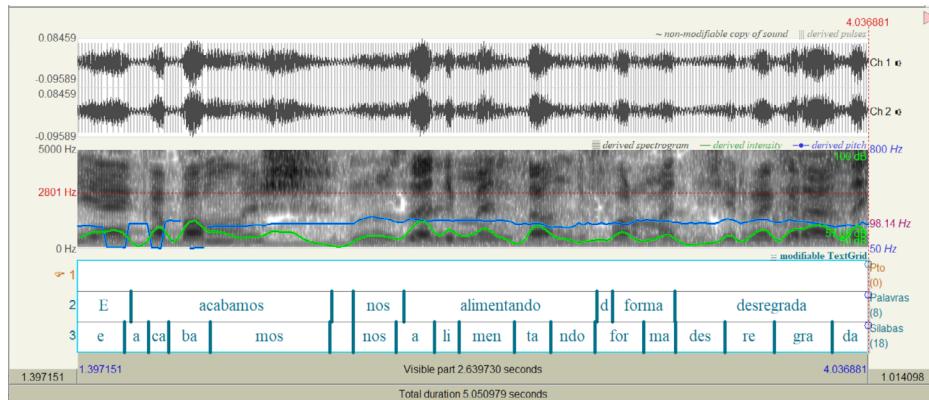
Ainda Vieira, F. (2016) explica que na variedade do PE se opta sistematicamente pela adjacência do pronome a v1, ou seja, “supõe-se que a ligação do pronome se efetive em relação a v1 e não a v2” (Vieira, F. 2016, p. 5). Nesse sentido, na variedade europeia, o clítico no interior do CV é produtivo com ênclise a v1, diferentemente da variedade brasileira na qual autores como Carneiro (2016); Vieira, F. (2016); Vieira e Vieira(2018); Araújo e Silva (2019), observam uma próclise generalizada a v2. Porém, nas variedades africanas do Português a problemática ganha mais expressão por terem a variedade europeia do português como norma-padrão.

Desta feita, antes de avançarmos com a apresentação e discussão dos dados, é importante comentarmos sobre alguns pontos que nos ajudam a nos posicionarmos sobre a ordem do clítico no interior do CV.

De acordo com Vieira, S. (2008, p. 6) alguns parâmetros acústicos (duração, intensidade) sugerem uma certa posição do clítico no interior do CV, pelo que a sua observação pode ser importante.

Nessa perspectiva, embora o estudo não tenha exclusividade de fazer uma análise acústica, fizemos algumas experimentações no Praat com foco na duração e intensidade da propagação das ondas sonoras. Abaixo temos demonstrações de algumas experimentações feitas.

Imagen 1 – Resultado de análise de alguns dados do corpus no Praat



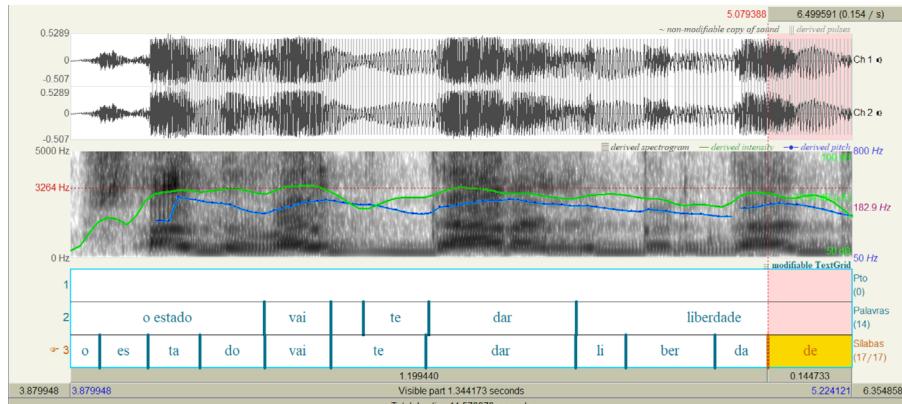
Fonte: elaboração própria.

Conforme a Figura 1, que traz a demonstração do comportamento sonoro e a visualização da fragmentação dos segmentos da fala “(4) *E acabamos nos alimentando de forma despregrada - Fi, I-FN*” no espectrograma, observamos que a propagação das ondas sonoras da articulação entre as palavras “acabamos” e “nos” está separada por uma pausa breve. A pausa é marcada pela linha fina que traduz o silêncio, que permite a separação entre o VAux e o clí-
tico, e a possível aproximação deste último à direita.

Ademais, observamos que em estruturas complexas em que ocorre a forma pronominal nos - como em (4) *E acabamos nos alimentando de forma despregrada e não prestamos atenção nos principais sintomas da diabetes que são esses que eu acabei de dizer, [...]* - Fi, I-FN - não é feita a supressão da última consoante fricativa, alveolar, - Voz [ʃ] presente no sufixo flexional (-mos), que ocorre em forma verbal flexionada na 1^a pessoa do plural (acabamos), o que levou a não colocação do hífen, sempre necessário com a elisão da última consoante do sufixo flexional (-mos) à luz da variedade do PE, ao anteceder o clítico. Tal cenário leva-nos a fundamentar a favor da ocorrência do clítico proclítico a v2.

As pausas entre o VAux e o clítico ocorrem em vários exemplos do nosso *corpus*, como no exemplo “(5) *o estado vai te dar liberdade [...]*”, conforme evidenciado na figura a seguir.

Imagen 2 – Resultado da experimentação da análise acústica no Praat



Fonte: Elaboração própria.

A Figura 2 mostra que há quase sempre na fala moçambicana uma pausa entre a articulação do VAux e o clítico no CV, marcada pela linha fina e menos preta, que traduz a inexistência da voz nessa parte, senão apenas ruído que não chega a atingir a intensidade de uma voz humana.

Outro dado não menos importante que observamos nos dados analisados, é que as estruturas complexas que apresentam a partícula “a” no interior (indicando uma ação contínua) e um verbo no infinitivo para formar o presente progressivo (equivalente ao gerúndio no PB), o clítico apresenta-se sempre hospedado a v2 ([6] *Então, o que eu estou a te dizer é: eu não estou a dizer que o meu país não tem problemas, meu país está cheio de problemas. Mas se tu jovens achares que a solução é apontar dedo a Frelimo, apontar dedo a Renamo, tu já falhaste* – Fii, I-BM).

Ademais, Mapasse (2005) atestou - na sua tese de doutorado, na qual trabalhou com colocação pronominal na variedade moçambicana do Português mobilizando a modalidade oral e escrita – que na modalidade escrita os participantes não fizeram a colocação do hífen no verbo auxiliar ligando o clítico, o que pode ratificar a ideia de este estar realmente hospedado a v2 e não a v1, conforme sucederia na norma-padrão.

5 Apresentação e discussão dos dados

Com a descrição e sistematização dos dados, confirmamos a observação de Vieira (2003), segundo a qual os complexos verbais apresentam poucas ocorrências em relação às lexias simples que são mais produtivas na fala cotidiana.

Assim, nos dados dessa pesquisa foi observado um total de 100 ocorrências de complexos verbais, formados por um ou dois verbos auxiliares. Sendo apenas três dados com 2 verbos auxiliares, todas em construções com verbo principal no infinitivo (*eu irei continuar a trabalhar, irei continuar a proporcionar-vos momentos de boa vibe* -Fii I-AM).

Nos complexos verbais com apenas duas formas verbais, os verbos auxiliares - uma vez sendo as formas funcionais - a sua flexão está distribuída entre o modo indicativo e subjuntivo, nos tempos de presente (indicativo), pretérito perfeito (indicativo) e pretérito imperfeito (subjuntivo), porém, observa-se também alguns no infinitivo (como em: *Posso sentar na minha casa e estar a dedicar-me especialmente ao programa Pé da Letra* – Fi, I-DM). E nos contextos com mais de um verbo auxiliar, o primeiro na ordem linear é que se encontra flexionado, e o segundo permanece no infinitivo (*o que estamos a tentar te dizer é que Moçambique não está totalmente independente*. Fi, I-DM).

Em função do verbo principal, das 100 ocorrências, foram observadas 2 (2%) com verbo principal no particípio, 2 (2%) no gerúndio e, com maior produtividade, 96 (96%) no infinitivo.

Nas duas ocorrências no particípio, o clítico está hospedado ao verbo auxiliar (uma na posição enclítica em contexto sem operador de próclise; e a outra na posição proclítica devido à ocorrência de uma categoria negativa), conforme demonstramos em (4) e (5):

- (7) *Esta mensagem que está sendo profanada, para mim é uma mensagem que não se pode sob ponto de vista nenhuma ser difundida. O que está acontecer é que nós temos crianças que estão a usar redes sociais.* (Fi, I-DM)

- (8) [...] mas por cada cliente **ter-se estar sentado** em casa e sem com que fazer, ou sem como iniciar o que a Renan quer fazer, vai colocar um emoji de riso. (Fii, I-AN)

Nas ocorrências com verbo principal gerundivo, o clítico está na posição intra-CV com próclise ao verbo principal ([4] *As diabetes tipo 2 são assintomática por se verificar na fase adulta e acabamos nos alimentando de forma desregrada* [...] Fi, I-FN).

Essas observações parecem confirmar a descrição de Mateus *et al.*, (2003, p. 25) segundo a qual “quando o verbo principal está no particípio ou no gerúndio, o clítico aparece obrigatoriamente proclítico ao verbo auxiliar/principal ou enclítico ao verbo auxiliar”.

Observando-se maior produtividade de complexos verbais com verbo principal ou temático infinitivo (96 = 96%), atentamos exclusivamente a elas por forma a compreendermos as variantes decorrentes em relação à ordem pronominal.

5.1 A ordem pronominal em complexos verbais com verbo temático infinitivo

A distribuição das ocorrências pelas duas cidades consideradas nesse artigo e pelas variantes posicionais relativas à ordem pronominal em complexos verbais, registra-se na Tabela 2:

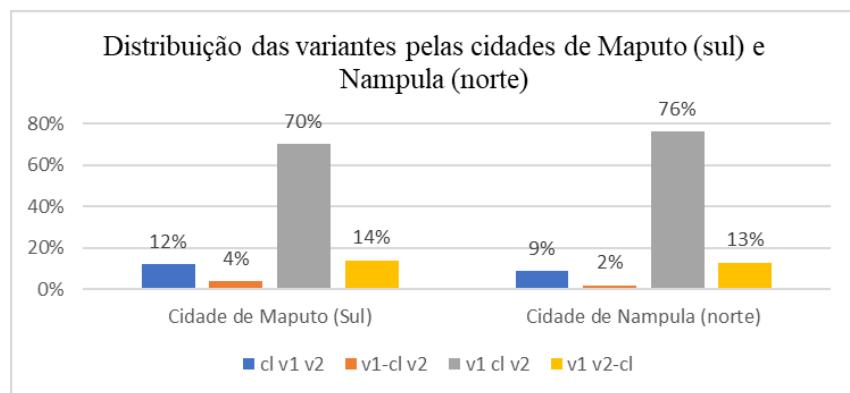
Tabela 2 – Distribuição dos padrões de colocação pronominal, nas duas cidades

	Cidade de Maputo (Sul)		Cidade de Nampula (Norte)		Total	
	Ocorrências	Percentual	Ocorrências	Percentual	10-10%	96-100%
cl v1 v2	6/51	12%	4/45	9%	10-10%	96-100%
v1-cl v2	2/51	4%	1/45	2%	3-3%	
v1 cl v2	36/51	70%	34/45	76%	70-73%	
v1 v2-cl	7/51	14%	6/45	13%	13-13%	

Fonte: Elaboração própria.

A distribuição da Tabela 2 permite-nos observar a variante mais produtiva ou preferencial nas duas cidades, e isso pode-se melhor visualizar através da representação gráfica a seguir:

Gráfico 1 – A ordem pronominal em complexos verbais nas duas cidades



Fonte: Elaboração própria.

Com os dados apresentados através da Tabela 2 ou Gráfico 1, fica nítido observar que a variante intra-CV com próclise ao verbo principal é a que mais ocorreu nas duas cidades, o que nos licencia a afirmarmos que esta constitui a variante/ordem pronominal (v1 cl v2), nos complexos verbais, produtiva e preferencial para os falantes considerados nas duas cidades (Maputo e Nampula). Segue a variante pós-CV (v1 v2-cl), depois a variante pré-CV (cl v1 v2) e, por último, a variante intra-CV com ênclide ao verbo auxiliar, uma variante que, segundo Vieira, F. (2016), é a mais produtiva na variedade europeia que até então vigora como padrão em Moçambique.

Destarte, de modo a compreendermos os fatores determinantes na ordem pronominal em complexos verbais nas duas cidades, assim como os contextos frásicos e morfossintáticos que favorecem/desfavorecem determinadas ordens, a seguir passamos a sistematizar os resultados focando no grupo dos fatores.

a) Tipos de clíticos

No âmbito dos estudos de ordem pronominal, Vieira, F. (2016) destaca a necessidade de se considerar as formas pronominais para a sistematização do comportamento das variáveis linguísticas. Assim, considerando essa variável, importa referir que no *corpus* analisado há predominância dos clíticos **me, te, se, nos** e **lhe**. E relativamente aos clíticos complementos **a** e **o**, não observamos nenhuma ocorrência.

Tabela 3 – Distribuição das três variantes segundo o tipo de clítico nas duas cidades – Complexos verbais

Cidade	Tipo de clítico								Total
	Ordem	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl
Me e nos	3-13%	2-8%	18-75%	1-4%	0	0	12-100%	0	36-37%
Te e lhe(s)	0	0	20-100%		1-20%	0	3-60%	1-20%	25-26%
Vos			2-50%	2-50%	0	0	0	0	4-4%
Se reflexivo	2-14%		9-64%	3-22%	2-13%	1-7%	7-47%	5-33%	29-30%
Formas contraídas	0	0	2-100%	0	0	0	0	0	2-2%
Total	5-5%	2-2%	51-53%	6-6%	3-3%	1-1%	23-24%	6-6%	96-100%

Fonte: Elaboração própria.

Com a tabela 3, é possível observarmos que a variante intra-CV com próclise ao verbo principal é a mais produtiva com todo tipo de clítico nas duas cidades, e ela parece ocorrer em todas as variáveis, desde as linguísticas até as sociais, como se pode observar nos exemplos:

- (9) *Pena que o meu telefone está tão partido que não posso vos mostrar. Mas eu falei com cunhado hoje ao celular logo que estas imagens me caíram. Eu liguei para o cunhado, e perguntei: cunhado, o que se passa, estamos preocupados, o que se passa? (Fii, I-AM)*
- (10) *Eu definitivamente, definitivamente. eu vou lhes dizer uma coisa, eu me vejo muito nos seus comentários, porque aqui nós não estamos a discutir Partidos, não estamos a discutir nenhuma, estamos a discutir propostas, e é exatamente isso. (Fi, I-DM)*
- (11) *[...] influenciamos as mulheres que devem se empoderar de uma maneira mais prática e simples, fazendo negócios, muitas outras coisas, ganharem independência financeira. (Fi, I-CN)*
- (12) *[...] tu vais ouvir por exemplo a Maria já é uma granda advogada em Nampula. Eu te pergunto... e olha que tu vais te perguntar, porque que a Maria aquela que veio de lá de baixo se tornou granda advogada. (Fi, I-AN)*

Segue a variante pós-CV (12%) em quase todas as formas clíticas quando juntadas com as ocorrências das duas cidades. As poucas ocorrências da variante pré-CV parecem depender de elementos proclisadores e associadas ao nível de escolaridade do informante. As formas contraídas estão em número de duas ocorrências (*nós dependemos de várias questões aqui, então não está sa afirmar categoricamente que Moçambique está independente* (Fi, I-DM)

Assim, com essas considerações surge a necessidade de compreender, através de valores absolutos e percentuais apresentados na tabela a seguir, o comportamento da colocação pronominal segundo a presença e ausência de elementos intervenientes no CV.

b) Presença e ausência de elementos intervenientes no complexo verbal

Esta constitui uma variável bastante determinante para a variação entre ênclise ou próclise na colocação dos clíticos pronominais nas sentenças.

Tabela 4 – Distribuição das três variantes segundo a presença e ausência de elementos intervenientes no CV

Presença e ausência de elementos intervenientes no CV										
Cidade	Cidade de Maputo (Sul)				Total	Cidade de Nampula (Norte)				Total
	Ordem	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl	
Presença	6-26%	1-4%	14-61%	2-9%	23/35	3-25%	0	7-58%	2-16%	12/35
Ausência	0	2-5%	30-75%	8-20%	40/61	1-5%	1-5%	14-66%	5-24%	21/61
Total	6-6%	3-3%	44-46%	10-10%	63/97	4-4%	1-1%	22-23%	7-7%	96-100%

Fonte: Elaboração própria.

Com a tabela 4, podemos observar que, nas duas cidades da variedade moçambicana, a variante pré-CV quase sempre depende da presença de elementos proclisadores, pois, na ausência deles, ela não é produtiva, tanto que nos dados da cidade de Maputo não houve sequer uma ocorrência em que o clítico estivesse na posição proclítica sem nenhum atrator; enquanto na cidade de Nampula apenas encontramos um dado nessas condições, o que julgamos ser pouco significativo. E isso é característico da variedade europeia, conforme lemos em Brito; Duarte e Mattos (2003) que a posição proclítica é induzida por fatores de natureza sintático-semântica ou prosódica (fazendo parte desses fatores, elementos como categoria negativa, conjunções, algumas preposições que funcionam como proclisadores).

Com maior produtividade lidera o ranking a variante “v1 cl v2” em ambas cidades, tanto em contextos com presença (22 – 33%), como com ausência (44 – 67%) de categorias proclisadoras. E essa corresponde a uma particularidade da variedade moçambicana em detrimento da norma europeia, pois nesta, a próclise ao V2 é condicionada pela presença de alguns elementos no interior do CV, como é o caso das preposições *de* e *por*, nas quais ocorre obrigatoriamente a posição proclítica ao verbo principal (Vieira e Vieira, 2018), e na ausência destes licencia-se a ênclise tanto no v1, como no v2, como também se permite a ênclise ao v2 na presença de atrator na posição pré-CV.

Assim, nos dados analisados e em contextos sem proclisadores, a ênclise ao V2 ocorre em um total de 13-21% e de 3-5% ao v1 de valores absolutos e percentuais, respectivamente. Assim, observa-se que a ênclise ao v1, neste âmbito, teve menos produtividade.

Do contrário, a produtividade da variante v1 cl v2, parece não dar relevância a variáveis como o tipo de oração, a distância entre V-CL ou CL-V, pois, em orações subordinadas em que estão presentes algumas conjunções com função de atratores na colocação dos clíticos, ela se mostra ainda efetiva, como também a existência de certos constituintes entre cl e v ou v e cl que traduz uma certa distância entre V-CL ou CL-V, parece não desfavorecer a variante v1 cl v2 nos dados analisados.

Relativamente às variáveis sociais, para além da região (cidade) que a sua distribuição numérica já vem sendo explicitada nas variáveis linguísticas discutidas neste texto, o que permitiu observar um comportamento semelhante em relação à efetividade de uma variante, sistematizamos os resultados através da Tabela 5 para explicar o comportamento da colocação pronominal segundo as variáveis de escolaridade, idade e gênero.

Tabela 5 – Distribuição das variantes segundo as variáveis sociais

Escalaridade					
Ordem	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl	Total
Ordem	cl v1 v2	v1-cl v2	v1 cl v2	v1 v2-cl	Total
Méd.	6-6%	0	25-78%	5-16%	32
Sup.	9-14%	3-5%	43-67%	9-14%	64
Total	11-11%	3-3%	69-71%	14-14%	96
Idade					
F-I	8-15%	1-2%	32-60%	12-23%	53
F-II	3-7%	2-5%	36-83%	2-5%	43
Total	11-11%	3-3%	69-71%	14-15%	96
Gênero					
H	7-10%	3-4%	50-72%	9-13%	69
M	3-11%	0	19-70%	5-19%	27
Total	10-10%	3-3%	70-72%	14-15%	96

Fonte: Elaboração própria.

De acordo com Beline (2010), “em falantes da mesma língua, a situação de fala ou registro, e ainda o nível socioeconômico do falante (escolaridade, socialização) podem ser determinantes na variação”.

E, de fato, ainda que com margens mínimas, nos falantes com nível de escolaridade superior, conforme a Tabela 5 apresenta, houve uma certa distribuição percentual maior entre as variantes em relação aos dados dos falantes com nível de escolaridade média, nos quais houve maior concentração de ocorrências (79%) na variante v1 cl v2. Tanto que, as poucas ocorrências de próclise a v1 devido a presença de elementos proclisadores ([13] a dificuldade que temos estado a enfrentar em Moçambique é uma verdade incontornável que não se pode tapar pela peneira – Fi, I-DM), foram observadas nas falas dos informantes com escolaridade superior, tal é o caso da fala que aqui nos serve de exemplo, que é formado em jornalismo, sendo ainda poeta, declamador e apresentador de programa televisivo sobre literatura moçambicana.

Todavia, tanto no nível médio, quanto superior registra-se a produtividade da variante intra-CV com próclise ao verbo principal, tal como decorre nas demais variáveis consideradas neste estudo.

No âmbito do estudo da variação, enquanto estágio que pode resultar em mudança, a idade permite compreender as tendências para uma mudança linguística, onde se atesta que quando ela mais se mostra na faixa menor que pode corresponder aos mais jovens, pode ser um forte indicador da tendência de mudança. Todavia, nas duas faixas consideradas nesse estudo, houve uma distribuição entre as quatro variantes, que alcança índices aproximados, e sempre a variante v1 cl v2 a mais produtiva, o que parece mostrar uma tendência de consolidação dessa ordem de colocação pronominal em complexos verbais na variedade moçambicana.

E o mesmo ocorre na distribuição das variantes segundo o gênero, isto é, houve, entre os homens e mulheres uma distribuição, nas quatro variantes, que alcança índices aproximados. Embora se costume dizer que as mulheres têm uma linguagem mais polida do que os homens e “usam menos formas estigmatizadas do que os homens e são mais sensíveis do que eles ao padrão de prestígio” (Labov, 1972c, p. 243 *apud* Tagliamonte 2012, p. 30), nos dados analisados houve uma distribuição equilibrada, tanto que na variante que tem se mostrado produtiva houve uma distribuição de 72% para os homens e 71% para as mulheres, um percentual bastante equilibrado.

Esse comportamento linguístico generalizado na colocação pronominal em CV pode ser explicado pela ausência de avaliações negativas em relação a essa variação em Moçambique. Sendo assim, o uso da variante intra-CV com próclise a V2 não é rejeitado, inclusive entre falantes com formação superior, o que reforça a sua aceitabilidade por parte de maior parte da população moçambicana. Além disso, a falta de familiaridade com os padrões formais de colocação em CV pensamos que pode contribuir para a ausência de autocensura, resultando nesses índices altos do clítico intra-CV.

Portanto, com a distribuição feita nas variáveis linguísticas e extralinguísticas/sociais controladas, comprehende-se que em todas, a variante intra-CV com próclise ao v2 (v1 cl v2) chama atenção pela sua produtividade, independentemente do contexto morfossintático ou dos elementos presentes na sentença, embora na variedade que vigora como padrão em Moçambique, essa ordem ocorra em contextos bem restritos como a presença de categorias proclisadoras no interior do CV.

Desta feita, essa produtividade, tanto de acordo com o tipo de clítico, tipo de oração, presença ou não do elemento proclisador: iv) distância entre V-CL ou CL-V ou tipo de estruturas verbais complexas, quanto de acordo com a região (cidade), sexo, idade e escolaridade dos participantes considerados neste texto, interpretamos como uma tendência da consolidação dessa variante variável na variedade moçambicana, transformando-se numa ordem pronominal com certo encaixamento na estrutura linguística e social, podendo estar a propiciar uma mudança na variedade moçambicana do Português.

6 Considerações finais

Nesse estudo, tínhamos o objetivo de analisar a variação da ordem pronominal em complexos verbais no Português falado em Moçambique, através do *corpus* constituído com base em falas extraídas de 3 vídeos de 3 programas televisivos da cidade de Maputo (Sul) e cidade de Nampula (Norte), onde foi crucial a sistematização dos dados, segundo determinadas variáveis linguísticas e sociais.

Quanto às variáveis linguísticas consideradas (tipo de clítico, tipo de oração, presença ou não do elemento proclisador, distância entre V-CL ou CL-V e tipo de estruturas verbais complexas), importa referir que houve uma distribuição entre as quatro variantes possíveis em complexos verbais (cl v1 v2, v1-cl v2, v1 cl v2, v1 v2-cl), no entanto, com maior produtividade da variante v1 cl v2.

O mesmo sucedeu quanto às variáveis sociais (cidade [região], sexo, idade e escolaridade). Porém, percebemos que os falantes com nível de escolaridade superior apresentam índices um pouco mais significativos de contextos de conformidade com a norma europeia,

quando comprados com os de escolaridade média, o que pode sugerir a ideia de que o nível de escolaridade contribui no freamento da variação, embora também tenham mostrado a produtividade da variante v1 cl v2.

Portanto, a variação da ordem pronominal em complexos verbais no Português falado em Moçambique que se reflete por meio da produtividade da variante intra-CV com próclise ao verbo temático infinitivo independentemente do contexto morfossintático, apresenta um grau de encaixamento bastante significativo na matriz linguística e social, que nos leva a pensarmos numa possível mudança em curso na variedade moçambicana em detrimento da norma europeia que vigora como padrão em Moçambique.

Declaração de autoria

A Contribuição dos autores foi: Bento Orlando Mutoba, na qualidade de autor principal, produção de dados, análise e interpretação dos mesmos, redação do texto e revisão; e Profa. Dra. Norma Lucia Fernandes de Almeida, na qualidade de coautora, discussão e decisões metodológicas, discussão dos dados e revisão do texto.

Agradecimentos

Na produção deste artigo, agradecemos à Profa. Dra. Huda Silva Santiago pelas discussões que suscitarão a ideia da produção deste artigo e pela avaliação do mesmo na disciplina de Linguística Histórica; à colega, Lara da Silva Cardoso pelas discussões iniciais sobre a temática; e ao nosso amigo, Arcedes José Manuel pela leitura final do texto.

Referências

- ARAÚJO, S. S. F.; SILVA, M. C. A. *A Sintaxe dos Pronomes clíticos no Português em Feira Santana-BA: Uma comparação com o Português luandense*, MACABÉA – REVISTA ELETRÔNICA DO NETLLI, Crato, v. 8., n. 2., p. 563-584, 2019. DOI: <https://doi.org/10.47295/mren.v8i2.1961>
- BELINE, R. A variação linguística. In: FIORIN, José Luiz. (org.). *Introdução à linguística: Objetos teóricos*. 6 ed. São Paulo: Contexto, 2022. p. 121-140.
- BRITO, A. M.; DUARTE, I.; MATOS, G. Tipologia e distribuição das expressões nominais. In: MATEUS, M. H. M. et al. *Gramática da Língua Portuguesa*. Tradução: 5 ed. Lisboa: Caminho, 2003. p. 795-867.
- CARNEIRO, Z. O. N. Colocação de clíticos em orações finitas em duas vertentes do português oral feirense: um contexto não variável. In: ALMEIDA, N. L. F. et al (Org.). *Variação Linguística em Feira de Santana – Bahia*. Feira de Santana: UEFS Editora, 2016, p. 141-174.
- GONÇALVES, P. A génese do português de Moçambique. Imprensa Nacional-Casa da Moeda, S.A, Lisboa, 2010.

- GONÇALVES, Perpétua. *Português de Moçambique: problemas e limites de padronização de uma variante não-nativa*. Universidade Eduardo Mondlane, Maputo, 2005.
- MAPASSE, E. L. A. *Clíticos pronominais em Português de Moçambique*. 2005. 166 f. Dissertação (Mestrado em linguística), Faculdade de Letras, Departamento de linguística Geral e Romântica, Universidade de Lisboa, 2005.
- MARTINS, A. M. A colocação dos pronomes clíticos em sincronia e diacronia. In: MARTINS, A. M.; CARRILHO, E. (org.). *Manual de linguística portuguesa*. Berlim/Boston: DeGruyter, 2016. p. 401-428.
- MATEUS, M. H. M. et alii. *Gramática da língua portuguesa*. 6ª edição. Lisboa: Editorial Caminho, 2003.
- MENDES, I. *Da neologia ao dicionário: o caso do Português de Moçambique*. Maputo, Texto Editores, 2010.
- MUTOBA, B. O.; ALMEIDA, N. L. F. *Análise variacionista do Português de Moçambique a partir da Música Kizomba: o caso de pronominalização “desviada” e ordem pronominal*. *Macabéa –Revista Eletrônica do Netlli*, Crato, v. 12, n. 4, p. 31-51, 2023. DOI: 10.47295/mren.v12i4.1254
- NGUNGA, A. Interferências de Línguas Moçambicanas em Português falado em Moçambique, 1ª ed. Centro Estudos Africanos, Universidade Eduardo Mondlane, Moçambique, 2012.
- NGUNGA, A. *Introdução à Linguística Bantu*, 2ª ed., Editor Imprensa Universitária (Universidade Eduardo Mondlane), Maputo, 2004.
- NHATUVE, D. *Reflexão sobre a normatização do português de Moçambique*, Universidade de Coimbra, Portugal, 2017.
- TAGLIAMONTE, S. Social Patterns. In: TAGLIAMONTE, S. *Variationist sociolinguistics: change, observation, interpretation*. Oxford: Wiley-Blackwell, 2012. p. 25-70.
- TIMBANE, A. A. *A variação linguística e o ensino do português em Moçambique*, *Confluência: revista do instituto da língua portuguesa*, N° 43 – 2º semestre de 2012, Rio de Janeiro, p. 261-2842012.
- TRINDADE, M. A. A. S. *Uma viagem sociolinguística pelas veredas do sistema de pronomes possessivos no português falado em luanda-angola*. 2021. 115 f. Dissertação (Mestrado em Estudos Linguísticos) - Departamento de Letras e Artes, Universidade Estadual de Feira de Santana, 2021.
- VIEIRA, M. F. *A ordem dos clíticos pronominais nas variedades urbanas europeia, brasileira e são-tomense: uma análise Sociolinguística do Português no início do século XXI*. 2016. 238f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2016.
- VIEIRA, M. F. *A cliticização pronominal em lexias verbais simples e em complexos verbais no português europeu oral contemporâneo: uma investigação sociolinguística*. 2011. 168 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2011.
- VIEIRA, S. R. *Análise de variedades do Português: a ordem dos clíticos em complexos verbais*. UFRJ. Texto apresentado no Congresso Internacional da ABRALIN, 2003 Disponível em <<https://www.catedraportugues.uem.mz/storage/app/media/bibliografia/silviarvabralin2003b.pdf>> Acessado aos 08 de fevereiro de 2024.
- VIEIRA, Silvia Rodrigues. *Colocação pronominal nas variedades europeia, brasileira e moçambicana: para a definição da natureza do clítico em Português*. 2002. 448 f. Tese (Doutorado em Língua Portuguesa) - Faculdade de Letras, Universidade Federal de Rio de Janeiro, 2002.

VIEIRA, S. R. *O complexo comportamento da ordem dos clíticos em complexos verbais*. CELSUL, Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2008.

VIEIRA, S. R.; CORRÊA, C. M. M. L. *Colocação pronominal no Português do Brasil*: a contribuição de estudos de percepção auditiva. *Letras de Hoje*, Porto Alegre, v. 52, p. 87-96, jan. 2017. DOI: [https://doi.org/10.15448/1984-7726.2017.1.25277](https://doi.org/https://doi.org/10.15448/1984-7726.2017.1.25277).

VIEIRA, S. R.; VIEIRA, M. F. A ordem dos Clíticos Pronominais no Português de São Tomé e Português de Moçambique. *In: BRANDÃO, S. Duas Variedades Africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Edgard Blücher Ltda, 2018. p.277-320.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. A língua como um Sistema diferenciado: Princípios empíricos para a teoria da mudança linguística. *In: WEINREICH, U.; LABOV, W.; HERZOG, M. Fundamentos Empíricos para uma Teoria de Mudança Linguística*. Tradução: Marcos Bagno. 1 ed. v. 4, São Paulo: Parábola Editorial, 2006. p. 87-125.

Demostración matemática: género discursivo y conexiones lógicas desde una mirada lingüística

Mathematical Proof: Discourse Genre and Logical Connexions from a Linguistic Perspective

Natalia Leiva Salum

Pontificia Universidad Católica de Chile
(PUC) | Santiago | CL
naleiva@uc.cl
<https://orcid.org/0009-0008-2538-7236>

Margarita Vidal Lizama

Pontificia Universidad Católica de Chile
(PUC) | Santiago | CL
mvvidal@uc.cl
<https://orcid.org/0000-0002-2161-8956>

Resumen: Una de las principales tareas que enfrentan quienes estudian Cálculo, Álgebra o Geometría en la formación terciaria es escribir lo que se conoce como una 'demostración matemática'. El propósito de una demostración es probar que una afirmación inicial es verdadera mediante un razonamiento deductivo que alterna el simbolismo matemático con el lenguaje natural. Más allá de encontrar el camino para resolver el problema, la escritura de este tipo de discurso multisemiótico encierra dificultades particulares; una de ellas es establecer relaciones lógicas, a través del lenguaje natural, que den sentido a las ecuaciones. El objetivo del presente estudio es ofrecer una descripción inicial de la demostración como género discursivo y los patrones de conexión que son relevantes en este. A partir de un trabajo interdisciplinario entre lingüistas y matemáticos, se caracterizó un conjunto de demostraciones confeccionadas por docentes de Licenciatura en Matemáticas de una universidad chilena. Dicha caracterización se fundamenta en el marco de la Lingüística Sistémico Funcional, particularmente en las descripciones de la semántica discursiva ideacional, conexión en español y el discurso matemático. Se plantea que la demostración posee tres etapas: Punto de la demostración \wedge Razonamiento matemático \wedge Confirmación. A su vez, se muestra que la demostración, al igual que géneros de otras disciplinas, despliega predominantemente conexiones causales externas realizadas congruente e incongruentemente y conexiones comparativas inter-



nas, realizadas congruentemente. Se concluye que la conexión en las demostraciones resulta crucial para formular ideacionalmente las proposiciones y para hilar las distintas etapas del género.

Palabras clave: Matemáticas; género; relaciones lógicas; conexión; Lingüística sistémico funcional.

Abstract: One of the main tasks faced by those studying Calculus, Algebra or Geometry at tertiary level is writing the text known as 'mathematical proof'. The purpose of a proof is to prove that an initial statement is true by deductive reasoning, through the alternation of mathematical symbolism and natural language. Beyond the solving of the problem itself, writing this kind of multi-semiotic discourse involves particular difficulties; one of them is to establish logical relationships, through natural language, that make sense of relations along the proof. The aim of this study is to offer an initial description of the proof as a genre as well as of the patterns of connexion that are relevant to it. Based on an interdisciplinary work between linguists and mathematicians, a set of proofs constructed by Mathematics professors at a Chilean university are described. This characterization is grounded on Systemic Functional Linguistics, with focus on ideational discourse semantics, connexion in Spanish and mathematical discourse. The paper proposes three stages for the proof: Point of the proof \wedge Mathematical reasoning \wedge Confirmation. The paper also shows that mathematical proof, like genres from other disciplines, displays predominantly external causal connexions realised congruently and incongruently, and internal connexions realised congruently. The analysis shows that connexions are crucial for the ideational formulation of propositions and for threading together the different stages of the mathematical proof genre.

Keywords: Mathematics; genre; logical relations; connexion; Systemic functional linguistics.

1 Introducción

La formación especializada en las disciplinas implica aprender no solo un cierto conocimiento, sino también una forma de crear y reproducir este conocimiento a través del lenguaje. Esta manera de entender la relación entre lenguaje y conocimiento puede aplicarse a todas las disciplinas, incluyendo aquellas que constituyen el dominio de las ciencias básicas. En este ámbito, generalmente la creación y transmisión de conocimiento se lleva a cabo a través de la integración de la semiosis verbal y otros sistemas semióticos, como representaciones visuales o lenguaje simbólico. Un caso interesante es el de las matemáticas, en el que una gran parte la construcción de conocimiento especializado pareciera asentarse en la interacción entre el simbolismo matemático – como una forma de semiosis – y el lenguaje natural. La integración de la semiosis verbal y el simbolismo matemático puede implicar un desafío importante en la formación de nuevos especialistas en el área, quienes tienen que enfrentarse a recursos semióticos complejos que realizan nuevos géneros, propios de la disciplina.

Desde hace ya varias décadas, existe un conjunto relevante de estudios, particularmente en inglés, que han abordado la matemática en su vínculo con el lenguaje y el aprendizaje. Estos estudios han apuntado a la estrecha relación entre habilidades lingüísticas y logro en matemáticas, indicando que el desarrollo de los conocimientos matemáticos depende en gran medida del desarrollo de una conciencia lingüística sobre cómo funciona el lenguaje en esta práctica. Esta orientación ha fundamentado investigaciones tanto en los niveles de formación escolar (Adams, 2003; Pimm, 1987) como en los años iniciales de formación universitaria (Cocking; Mestre, 1988). En los trabajos mencionados se observa una amplia conciencia sobre la circulación de diferentes textos en el ámbito de esta práctica, tales como las presentaciones del profesor, las explicaciones de los textos, los problemas y los exámenes.

Uno de los textos más relevantes de la práctica de las matemáticas en los primeros años de formación especializada es la resolución de problemas, también conocida como ‘demostración matemática’. Una demostración matemática, desde una mirada disciplinar, puede ser descrita en términos generales como la prueba de un conocimiento matemático. Esta prueba involucra un punto de partida establecido *a priori*, a partir del cual se desarrolla un proceso de deducción lógica que culmina en la comprobación (o no) del conocimiento en cuestión (Alfaro-Carvajal *et al.*, 2019; Alvarado & González, 2009; Bustos & Zubietta, 2019; Fiallo *et al.*, 2013; Martínez, 2022). Si bien la demostración matemática ha sido foco de múltiples estudios, no pareciera existir una definición compartida ni una caracterización común de este género por parte de los especialistas de la disciplina (Fiallo *et al.*, 2013; Lew; Mejía-Ramos, 2020); más aún, se identifican en la literatura diversos sentidos, categorías, funciones y contextos de uso atribuidos a la demostración matemática (Alfaro-Carvajal *et al.*, 2019). Por otra parte, desde una mirada centrada en la enseñanza, manuales (e.g. Sundstrom, 2021) o apuntes no publicados escritos por matemáticos (e.g. Allahbakhshi *et al.*, 2022; Wilson, s/a) ofrecen orientaciones para la escritura de demostraciones que sugieren algunos de sus posibles rasgos lingüísticos, como una evidente organización lógica, el uso de términos y conceptos técnicos, mediante lenguaje natural y lenguaje matemático, y la definición de símbolos y notación utilizados en el transcurso de la demostración (Lew & Mejía-Ramos, 2020).

Un cuerpo importante de investigación sobre la demostración matemática se desarrolla en el ámbito de la enseñanza de las ciencias, generalmente en el contexto de educación

secundaria. En este ámbito se observa un interés sostenido por identificar las dificultades que enfrentan las y los matemáticos en formación en el aprendizaje de la demostración, así como por presentar estrategias didácticas que los apoyen en el proceso de aprendizaje (Bustos & Zubieta, 2019; Camacho *et al*, 2014; Martínez, 2001; Sua Flores, 2019). Las investigaciones señalan como dificultades aspectos diversos, entre los que se cuentan la complejidad de poner en práctica el razonamiento lógico adecuado para el desarrollo de una demostración y el uso inadecuado de casos para exemplificar axiomas (Alvarado & González, 2009; Camacho *et al*, 2014; Fiallo *et al*, 2013). Otro conjunto de investigaciones vincula la demostración matemática con la práctica de la argumentación, considerando generalmente el modelo argumentativo de Toulmin (e.g. Katz *et al*, 2023; Urhan & Zengin, 2024).

En la investigación de orientación lingüística, es posible identificar algunos estudios que abordan la matemática para describir sus rasgos característicos. En inglés, por ejemplo, se identifica el trabajo de Spanos *et al* (1988), quienes describen los rasgos lingüísticos que se ponen en juego en la resolución de problemas matemáticos en la escuela, en los niveles sintáctico, semántico y pragmático, señalando su alto grado de sofisticación. Según estos autores, entre los rasgos de nivel sintáctico que caracterizan el registro matemático y que pueden ser desafiantes para los aprendientes, destacan el uso de conectores lógicos (e.g. *if...then; given that...*). Más recientemente, Lew y Mejía-Ramos (2020) describen algunas convenciones lingüísticas para la escritura de demostraciones matemáticas en el ámbito de educación terciaria, teniendo en cuenta diferentes contextos de uso (textos de estudio, demostraciones producidas por estudiantes, demostraciones producidas por profesores en la pizarra). Si bien esta investigación plantea la relevancia de entender la demostración como un género disciplinar propio de las matemáticas, integrando una mirada sobre el lenguaje natural, su descripción emerge de las representaciones de los docentes sobre 'infracciones' (*breaches*) o errores en el proceso de su escritura, y no de la descripción de patrones lingüísticos en este género.

Un conjunto de investigaciones que ha abordado el discurso de la matemática desde una mirada lingüística se fundamenta en la Lingüística sistémico-funcional (LSF). En este dominio, el trabajo fundacional y más extenso sobre la matemática es el de O'Halloran (1999; 2005), quien la ha explorado como un discurso esencialmente multisemiótico, asumiendo una mirada gramatical hacia las imágenes visuales y el simbolismo en las matemáticas. Otros trabajos desde la LSF se han enfocado en los desafíos pedagógicos del discurso matemático a nivel escolar para, a partir de esto, proponer estrategias que apoyen el aprendizaje de la matemática en los estudiantes (Accurso *et al*, 2017; O'Halloran, 2015; Schleppegrell, 2007; Segerby, 2017). Más recientemente, algunas investigaciones han abordado la matemática y su rol en la física, particularmente en manuales de nivel secundario en inglés. Estos estudios describen la organización textual en el lenguaje matemático, tal y como se despliega en géneros propios de la física (Doran, 2018a, b), reflexionando sobre la dimensión semiótica del lenguaje matemático (Doran, 2022). La aproximación al discurso de la matemática desde esta perspectiva ha considerado fundamentalmente textos que circulan en la educación secundaria. Estos estudios han contribuido descripciones relevantes sobre el discurso matemático en inglés desde una mirada sistémico funcional.

Considerando los estudios hasta ahora revisados, es posible señalar que la investigación lingüística sobre el discurso matemático se ha desarrollado en particular en el contexto de habla inglesa, mientras que en español este objeto se ha abordado casi exclusivamente desde el ámbito de la enseñanza de las ciencias. Teniendo en cuenta la aparente ausencia

de estudios lingüísticos sobre el discurso matemático, este artículo plantea dos objetivos generales. Primero, ofrecer una descripción inicial de la demostración matemática como un género relevante del discurso matemático, desde una mirada lingüística especializada, que complemente las propuestas disponibles desde la didáctica de la matemática. Segundo, describir las relaciones lógicas que permiten elaborar el razonamiento deductivo en este género en español. En este sentido, el artículo contribuye una descripción lingüística de este objeto, que no ha sido abordado hasta ahora desde esta perspectiva. El foco en las relaciones lógicas en la demostración responde a la importancia del razonamiento lógico en el despliegue de este género, cuestión que se evidencia en un número importante de investigaciones sobre este objeto (e.g. Schleppegrell, 2007; Sundstrom, 2021; Lew & Mejía-Ramos, 2020; Martínez, 2022, entre otros). Para abordar la exploración del género demostración matemática y de las relaciones lógicas en él, este estudio se fundamenta en la Lingüística sistémico funcional (Halliday & Matthiessen, 2014; Martin, 1992). Esta aproximación a la demostración matemática busca contribuir conocimiento lingüístico relevante que pueda ser aplicado en la formación de las y los estudiantes en matemáticas, para quienes la escritura de la demostración es un importante desafío.

El artículo se organiza de la siguiente forma. Se introducen a continuación los fundamentos teóricos sobre los que se sustenta el estudio de la demostración matemática, teniendo en cuenta principios generales de teoría sistémica, la perspectiva teórico-metodológica asumida en esta descripción y los sistemas de significados relevantes para ella. Luego, se describen los aspectos metodológicos del estudio, con foco en los procedimientos analíticos. A continuación, se presentan los resultados, que inician con una descripción de la estructura esquemática propuesta para el género demostración matemática, para introducir luego los patrones de CONEXIÓN observados en los datos. Finalmente, se proponen las conclusiones más relevantes del trabajo y se señalan algunas proyecciones para el estudio de la demostración matemática en el ámbito de educación terciaria.

2 Fundamentos teóricos

En esta sección se presentan los fundamentos en que se basa el estudio propuesto. Específicamente, se abordan los principios generales de la teoría sistémica (2.1), con énfasis en aquellos conceptos relevantes que permiten describir la perspectiva trinocular asumida en este estudio. A continuación, se profundiza en el estrato del registro, particularmente en la variable campo, y se presentan las opciones de significado ideacional clave en el estrato semántico-discursivo (2.2). Luego, se introducen los sistemas de CONEXIÓN (2.3) y de PERIODICIDAD (2.4).

2.1 Aproximación teórica. Principios generales de teoría sistémica

La Lingüística sistémico funcional (en adelante, LSF) es una teoría general sobre el lenguaje. En esta teoría, se entiende el lenguaje como una herramienta para crear significados en la vida social (Halliday & Matthiessen, 2014), es decir, como un sistema semiótico que funciona siempre en relación con el contexto. En este marco, el contexto se conceptualiza como

un sistema semiótico más abstracto que el lenguaje. En la propuesta de contexto de Martin (1992) aquí asumida, se identifican dos niveles diferentes en el contexto, género y registro. El género ha sido definido, desde un enfoque pedagógico, como un proceso social orientado a un propósito y dividido en etapas (Martin & Rose, 2008). Desde una perspectiva que enfatiza la dimensión lingüística, género se entiende como una configuración recurrente de significados textuales, ideacionales e interpersonales, orientada a lograr un propósito cultural determinado (Martin & Rose, 2007). Estas configuraciones de significado están organizadas en etapas, que contribuyen a lograr el propósito global del género. Por ejemplo, el género 'narración' se organiza en las siguientes etapas:

Orientación \wedge Complicación \wedge Resolución \wedge Coda

La secuencialidad de cada una de las etapas se señala por el símbolo \wedge que aparece entre cada una de ellas. El nombre de cada etapa se escribe con mayúscula inicial, debido a que se trata de una función en términos sistémicos. Las etapas que no están entre paréntesis corresponden a etapas obligatorias en la estructura esquemática, mientras que las etapas opcionales se escriben entre paréntesis.

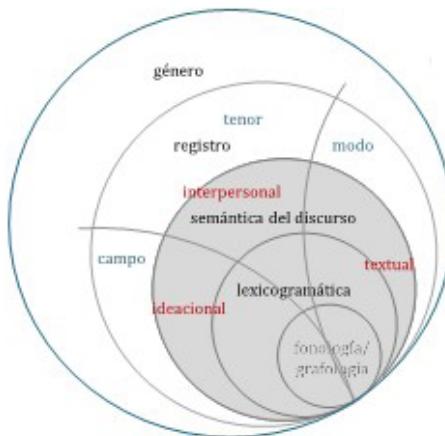
El siguiente nivel del contexto, denominado registro, distingue las dimensiones del contexto inmediato de un texto que influyen en los patrones de significado que lo constituyen. Estas dimensiones corresponden a las variables de campo, tenor y modo. Por ejemplo, en una narración el campo suele ser del ámbito del sentido común, el tenor suele implicar una relación de relativa cercanía entre los hablantes y el modo suele ser monologal, tanto oral como escrito (Eggins, 2004). La conceptualización del contexto, en la LSF, es fundamental, en la medida que guía la interpretación de los fenómenos lingüísticos estudiados.

La arquitectura teórica de la LSF se organiza alrededor de un conjunto de principios teóricos generales. En esta investigación, tres de estos principios son clave para fundamentar la aproximación a las relaciones lógicas en la demostración matemática: estratificación, realización y metafunción. El principio de **estratificación** permite identificar niveles o estratos tanto en el contexto como en el lenguaje. En este último se distinguen tres estratos diferentes: fonológico-grafológico, léxico-gramatical y semántico-discursivo (Halliday & Matthiessen, 2014). En la LSF, se entiende que entre los estratos del lenguaje y el contexto se establece una relación de **realización**, tal como entre el contexto y el lenguaje: los significados en el nivel superior son codificados o simbolizados por los significados del nivel inferior (Martin & Rose, 2007). Así, los significados semántico-discursivos son realizados a través de patrones léxico-gramaticales, tal y como estos últimos son realizados por patrones fonológicos-grafológicos.

El principio **metafuncional** supone que el lenguaje cumple tres grandes funciones en la vida social: construir la experiencia interna y externa a través de la semiosis; establecer relaciones sociales entre los usuarios del lenguaje en diversos contextos; y organizar estos recursos de significado en la forma de textos coherentes. Estas funciones, denominadas metafunciones, son la metafunción ideacional, la metafunción interpersonal y la metafunción textual, respectivamente. El modelo de contexto propuesto en el marco de la LSF asume de manera general que estas metafunciones se proyectan en el nivel del registro, de modo que cada una de ellas se relaciona con una de las variables registrales: la metafunción ideacional se relaciona con la variable de campo, la metafunción textual con la variable de modo y la metafunción interpersonal con el tenor (Halliday & Matthiessen, 2014; Martin, 1992).

La Figura 1 representa estos tres principios centrales en la LSF. Los diferentes estratos en el lenguaje y el contexto se representan en círculos concéntricos; la relación de realización es representada mediante una línea recta que atraviesa los estratos y las metafunciones se proyectan sobre el lenguaje y el registro en el contexto.

Figura 1 – Principios teóricos en LSF



Fuente: adaptado de Martin & Rose (2008, p.17)

Esta conceptualización teórica sobre el lenguaje permite abordar el estudio de los fenómenos lingüísticos desde una mirada teórico-metodológica que se ha denominado **trinocular** (Halliday & Matthiessen, 2014). La perspectiva trinocular se sustenta en los principios de estratificación, realización y metafunción descritos previamente, y supone explorar un fenómeno a partir de tres puntos de vista complementarios, de manera simultánea: ‘por arriba’, es decir, considerando el estrato superior al estrato en que se identifica el fenómeno; ‘por alrededor’, es decir considerando patrones de significado complementarios en el mismo estrato; y ‘por abajo’, es decir, teniendo en cuenta la realización del significado en el estrato inmediatamente inferior (Halliday & Matthiessen, 2014). Esta aproximación ha sido aplicada al estudio de diversos fenómenos lingüísticos, en diferentes estratos del lenguaje (e.g. Hao, 2020; Quiroz & Martin, 2021). De esta manera, para el estudio de la CONEXIÓN en las demostraciones matemáticas, la mirada trinocular implica considerar, ‘por arriba’, el nivel del registro; ‘por alrededor’, otros tipos de significados en el nivel semántico-discursivo y, ‘por abajo’, las realizaciones léxico-gramaticales en español. Esta perspectiva permite proveer una descripción detallada y teóricamente sustentada de la CONEXIÓN en las demostraciones matemáticas.

2.2 La perspectiva trinocular: del registro a la semántica-discursiva

Dentro del marco de la LSF, el punto de entrada más productivo para el estudio del lenguaje disciplinar suele ser la metafunción ideacional. Mirado ‘por arriba’, esto implica considerar la variable campo en el registro. El **campo** se entiende en términos técnicos como un recurso semiótico para la construcción de la experiencia desde una perspectiva estática o dinámica (Doran & Martin, 2021). Mientras la perspectiva estática implica un foco en los **ítems** y sus rela-

ciones, la perspectiva dinámica implica un foco en las **actividades**. Martin (2020) establece una distinción entre **actividad** (unmomented activity) y **serie de actividades** (momented activity)¹.

En el estrato semántico-discursivo, la unidad que realiza una actividad del campo es la **figura**. Una figura se entiende como una configuración semántico-discursiva de un estado o una ocurrencia. Esta configuración se organiza alrededor de una entidad o dos entidades relacionadas, en el caso de las figuras de estado, y de una ocurrencia, en el caso de las figuras de ocurrencia². En el estrato léxico-gramatical, la unidad que realiza la figura, de manera natural o congruente³, es la cláusula. Las Tablas 1 y 2 a continuación introducen ejemplos de figuras de estado y de figura de ocurrencia, respectivamente, con su realización congruente en la léxico-gramática.

Tabla 1 – Figura de estado: ‘Tres es un número primo’

semántica-discursiva	figura de estado		
	entidad	=	entidad
	Tres	es	un número primo
léxico- Gramática	Participante grupo nominal	Proceso grupo verbal	Participante grupo nominal
cláusula			

Fuente: elaboración propia

Tabla 2 – Figura de ocurrencia: ‘Los estudiantes aprendieron sobre los números primos en la escuela’

semántica-discursiva	figura de ocurrencia			
	entidad	ocurrencia	entidad	entidad
	Los estudiantes	aprendieron	sobre los números primos	en la escuela
léxico-gramática	Participante grupo nominal	Proceso grupo verbal	Participante frase preposicional	Circunstancia frase preposicional
cláusula				

Fuente: elaboración propia

¹ Seguimos aquí la traducción ofrecida en Leiva (2022)

² Para ver una explicación detallada de la estructura orbital de las figuras, ver Hao (2020).

³ La gramática que se encuentra en una relación natural o directa con la semántica se conoce como ‘congruente’ (Halliday 1994): los procesos son realizados como grupos verbales, los participantes como grupos nominales, las circunstancias como adverbios o grupos preposicionales, las cualidades como adjetivos y las conexiones como conectores.

Dos o más figuras pueden conectarse en una **secuencia**, lo que se realiza congruentemente en un **complejo clausal** en el estrato léxico-gramatical. 'Por arriba', una secuencia de figuras realiza una serie de actividades en el campo. Las correlaciones no marcadas entre unidades ideacionales en la léxico-gramática, la semántica-discursiva y el registro se presentan en la Tabla 3.

Tabla 3 – Correlaciones ideacionales no marcadas

registro	semántica-discursiva	léxico-gramática
serie de actividades	secuencia	complejo clausal
actividad	figura	cláusula
ítem	entidad	grupo nominal

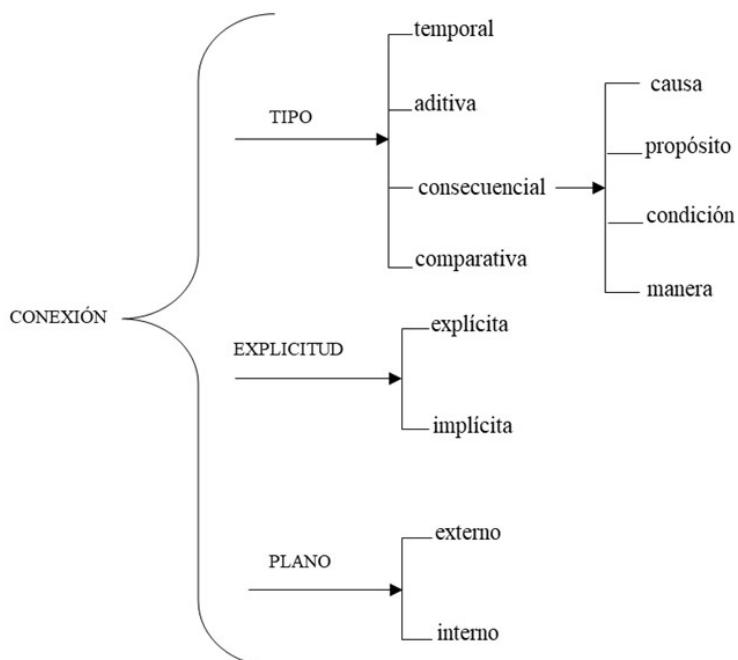
Fuente: Martin, 2020.

Desde un punto de vista semántico-discursivo, los significados que funcionan para establecer relaciones lógicas entre figuras, es decir, para formar secuencias, se organizan en el sistema de **CONEXIÓN**, que se presenta a continuación.

2.3 Sistema de CONEXIÓN

El sistema de **CONEXIÓN** (Martin, 1992; Hao, 2020) organiza los significados ideacionales de naturaleza lógica en el estrato semántico-discursivo. Estos significados funcionan para relacionar o conectar sucesos o estados que se construyen como figuras. Los significados lógicos se realizan de manera prototípica como conectores. En el marco de la LSF, el sistema de **CONEXIÓN** es modelado como una red compuesta por tres sistemas de opciones que funcionan de manera simultánea, como se ilustra en la Figura 2.

Figura 2 – Sistema de CONEXIÓN



Fuente: Hao, 2020; Martin, 1992

Uno de estos sistemas organiza las opciones relativas al tipo de conexión, correspondientes, en un primer nivel de delicadeza, a conexiones temporales, aditivas, consecuenciales y comparativas. Un segundo sistema distingue entre conexión explícita y conexión implícita, es decir, realizada o no por un ítem léxico. Por último, un tercer sistema distingue el plano de la conexión, es decir, entre conexión externa y conexión interna. La primera está orientada a la representación experiencial del significado en el texto y, por tanto, al despliegue de actividades en el campo; la segunda está orientada a la organización retórica del texto y, por lo tanto, al género (Halliday & Hasan, 1976; Martin, 1992, p. 180).

Los criterios para distinguir entre conexión externa e interna no han sido extensamente elaborados en la propuesta de Martin (1992). Sin embargo, se proponen ciertas consideraciones respecto de la conexión interna que ayudan a distinguirla de la externa. Por un lado, desde un punto de vista textual, la conexión interna tiende a ser cohesiva, i.e. a funcionar entre complejos clausulares o segmentos mayores de texto (Martin, 1992; Martin & Rose, 2007). Por otro lado, desde una mirada léxico-gramatical, es posible parafrasear una conexión interna como una conexión externa añadiendo además un proceso verbal o mental que proyecte una de las figuras conectadas; de esta manera se modifica además la taxis del complejo clausal, es decir, de no-hipotáctica a hipotáctica o viceversa (Hao, 2020, p. 111; Martin, 1992, p. 226).

Al igual como otros significados semántico-discursivos, la conexión puede realizarse de manera congruente o incongruente en la léxico-gramática. Una realización congruente se da cuando los significados discursivos se relacionan de manera natural o directa con su realización en la léxico-gramática: las ocurrencias son realizadas como Procesos, las entidades como Participantes, Cualidades y/o Circunstancias, y las conexiones como conectores (Hao, 2020; cf. Halliday & Martin, 1993). De este modo, la realización congruente de una secuencia de figuras es en general un complejo clausal, como se señaló más arriba. Un ejemplo de realización congruente de secuencia, con las figuras realizadas por cláusulas y la conexión realizada por un conector, se presenta a continuación:

Tabla 4 – Realización congruente de una secuencia de figuras

semántica-discursiva	secuencia		
	figura	conexión	figura
	Margarita estaba preocupada	YA QUE	su itinerario cambió
léxico-gramática	cláusula	conector	cláusula
		complejo clausal	

Fuente: elaboración propia

Además de su realización congruente, una secuencia puede realizarse de manera incongruente o metafórica a través de una cláusula. Esto implica que la conexión lógica es realizada por una estructura diferente de la prototípica, como un grupo verbal, un grupo nominal, un adjetivo o una frase preposicional (Coffin, 2004; Martin, 2004). Se ha descrito que esta realización incongruente implica una tensión entre el estrato de la léxico-gramática y la semántica-discursiva (Halliday & Mathiessen, 2014; Martin, 2020). Para que una secuencia se realice metafóricamente en una cláusula es necesario que al menos una de las figuras

que la compone se realice en el rango del grupo o la frase (Martin, 1992, pp. 169-170), es decir, que sea ella misma una realización metafórica o incongruente de una figura. Un ejemplo de realización incongruente de secuencia se presenta a continuación:

Tabla 5 –Realización incongruente de secuencia mediante nominalizaciones

semántica-discursiva	secuencia		
	figura	conexión	figura
	El cambio en su itinerario	PRODUJO	la preocupación de Margarita
léxico-gramática	cláusula		

Fuente: elaboración propia

En la Tabla 5, ambas figuras son realizadas en el rango del grupo mediante nominalizaciones ('preocupación' y 'cambio'), las que son conectadas mediante el grupo verbal 'produjo'. Si bien la forma canónica de realizar metafóricamente una figura es la nominalización (Martin, 1993), también puede producirse mediante una referencia textual o una cláusula incrustada (Christie & Derewianka, 2008, p. 26; Halliday & Matthiessen, 2014, p. 717; Hao, 2020, p. 125; Martin, 1992, p. 140). La Tabla 5 a continuación muestra una secuencia cuya primera figura es realizada en el rango del grupo, mediante el ítem de referencia textual 'esto'. Este demostrativo neutro tiene la capacidad de recuperar anafóricamente una figura introducida previamente en el discurso:

Tabla 6 – Realización incongruente de secuencia mediante referencia textual

semántica-discursiva	secuencia		
	figura	conexión	figura
	Esto	PRODUJO	la preocupación de Margarita
léxico-gramática	cláusula		

Fuente: elaboración propia

En la secuencia expuesta en la Tabla 7, la primera figura es realizada en el rango del grupo, mediante una cláusula incrustada, la que se indica con doble paréntesis cuadrados:

Tabla 7 – Realización incongruente de secuencia mediante cláusulas incrustadas

semántica-discursiva	secuencia		
	figura	conexión	figura
	[[Que el itinerario cambiara]]	PRODUJO	la preocupación de Margarita
léxico-gramática	cláusula		

Fuente: elaboración propia

El estudio de la CONEXIÓN en la demostración matemática es un aspecto central de este artículo. En el marco de la LSF, el trabajo más relevante a la fecha sobre significados lógicos en matemáticas es el de O'Halloran (2005). La autora propone que en las matemáticas las conexiones se realizan tanto en el simbolismo matemático como en el lenguaje; las conexiones lingüísticas, plantea, suelen funcionar más en la organización retórica del texto (plano interno) que en la construcción de la experiencia (plano externo). Sin embargo, su descripción no ofrece criterios explícitos para distinguir entre estos dos planos ni presenta un análisis que diferencie unidades en los estratos léxico-gramatical y semántico-discursivo.

2.4 Sistema de PERIODICIDAD, una mirada 'por alrededor'

La perspectiva trinocular que fundamenta esta exploración considera 'por alrededor' significados textuales en el estrato semántico-discursivo. En particular, el estudio de los recursos de CONEXIÓN se complementa con una consideración del despliegue de recursos textuales propios del sistema de PERIODICIDAD (Martin & Rose, 2007). Este sistema aborda los significados textuales que permiten organizar el texto como un cúmulo de información. Se distinguen dos puntos de prominencia posibles en el texto, una prominencia temática –que ancla un punto de partida para un cúmulo de información – y una prominencia informativa –que presenta aquellos significados propuestos como lo 'nuevo'. Estos dos puntos de prominencia pueden observarse en distintos niveles.

A nivel de sección o párrafo de un texto, los dos puntos de prominencia se denominan hiperTema e hiperNuevo. Se reconocen como hiperTema los significados ideacionales e interpersonales que predicen la información desplegada en un párrafo y que aparecen en posición inicial de este, generalmente en lo que se conoce como oración tópico. En tanto, el hiperNuevo corresponde a significados ubicados en posición final, orientados a acumular la información desplegada en la sección o párrafo del texto (Martin 1992; Martin & Rose, 2007). Los hiperTemas suelen presentarse en los textos de manera más recurrente que los hiperNuevos, es decir, los textos suelen más recurrentemente predecir que consolidar información. Esta organización periódica de la información en los textos no ocurre de forma natural, sino que depende de los recursos que pueda desplegar el hablante para organizar la información en el texto. En otras palabras, no todos los textos ni todos los párrafos de un texto se organizan necesariamente a partir de hiperTemas e hiperNuevos. El siguiente fragmento exemplifica la predicción de información en un párrafo mediante un hiperTema, destacado en negrita:

El día viernes se llevó a cabo el Festival escolar de Matemáticas. Tras una larga organización, profesores de toda la comuna dieron inicio al evento que busca familiarizar a los estudiantes con esta temida asignatura. Los estudiantes aprendieron sobre los números primos, ejercitaron operaciones básicas y tuvieron sus primeros acercamientos a las ecuaciones. El festival concluyó con una competencia de cálculo mental, tras la cual se premió a los estudiantes ganadores.

Los momentos de predicción (i.e. Temas) y de acumulación de información (i.e. Nuevos) pueden identificarse por dos patrones clave. Desde un punto de vista ideacional, en estos momentos se introducen entidades de carácter más general que luego se elaboran en el desarrollo informativo. Desde un punto de vista interpersonal, se tienden a negociar patrones

de valoración de manera explícita. Estos patrones ofrecen criterios para determinar la presencia de estos puntos de predicción y acumulación de la información en el despliegue textual.

La cláusula también puede entenderse, desde una mirada textual, como un cúmulo de información en el que se distribuyen la prominencia temática y la prominencia informativa. En este nivel, estas se denominan Tema y Nuevo, respectivamente. En términos generales, se ha identificado como el Tema de la cláusula en inglés todos los elementos que anteceden al Proceso, mientras que el Nuevo se ha descrito como todos los elementos que lo suceden (Halliday & Matthiessen, 2014). Se distinguen, además, diferentes tipos de Tema en la cláusula, dependiendo del tipo de significado que se realiza: Tema textual, Tema interpersonal, Tema tópico marcado y Tema tópico no marcado⁴. Las opciones de Tema y Nuevo se ilustran en la Tabla 7, exceptuando el Tema interpersonal, debido a que el componente de significado interpersonal ha sido señalado como poco relevante en el discurso de las matemáticas (Doran, 2020). Se ejemplifican en la Tabla 8 dos opciones de distribución de la información en la cláusula, de modo de sugerir las diferentes posibilidades de organización de la información.

Tabla 8 – Ejemplos: análisis de Tema y Nuevo en la cláusula

Tema textual	Tema tópico marcado	Tema tópico no marcado	Proceso	Nuevo
De esta manera,	durante este año	los estudiantes	aprendieron	sobre los números primos.
		Los estudiantes	aprendieron	este año sobre los números primos.

Fuente: elaboración propia

Desde una perspectiva multisemiotica de la demostración matemática, Doran (2018b) y O'Halloran (2005) han propuesto métodos para el análisis del Tema en la simbología matemática, es decir, considerando la distribución de la información del lenguaje no verbal o simbólico. Sin embargo, en este estudio el análisis de la organización de la información en la demostración matemática se enfocará en las cláusulas, dentro de las cuales a menudo se inserta el simbolismo matemático. Por este motivo, se tendrá en cuenta los criterios generales de Moyano (2010; 2021) para el análisis de Tema y Nuevo en español. La autora propone que en esta lengua el Tema (no marcado) corresponde al participante que concuerda en persona y número con la desinencia del Proceso, mientras que el Tema marcado se realiza mediante elementos circunstanciales en posición inicial de cláusula y el Tema textual a través los elementos conectivos en posición inicial de cláusula.

En esta sección se han presentado los fundamentos teóricos a partir de los cuales se explora lingüísticamente la demostración matemática. Estos fundamentos informan los criterios y principios metodológicos del estudio, particularmente respecto del análisis de la demostración como un género y de las opciones de CONEXIÓN que en él se despliegan.

⁴ Ver Halliday & Matthiessen (2014). 3.4 'Textual, interpersonal and topical Themes' (pp. 105-114).

3 Aspectos metodológicos

Esta investigación es un estudio de caso exploratorio de carácter descriptivo y cualitativo. El propósito es abordar de manera profunda, a partir de un análisis metafuncionalmente diferenciado y de la mirada trinocular de la LSF, la descripción del género y los patrones de CONEXIÓN en un grupo de instancias de la demostración matemática. A continuación, se detallan los datos seleccionados y los procedimientos de análisis empleados para plantear los resultados.

3.1 Datos

Los datos seleccionados para este estudio corresponden a 5 demostraciones matemáticas escritas por docentes del curso Taller de Matemáticas, de primer año del currículum de Licenciatura en Matemáticas en una universidad tradicional de Chile. Estas demostraciones fueron escritas en el marco de un proceso de intervención para la integración de habilidades de comunicación escrita en el currículo de Matemáticas. Por ello, las demostraciones que componen el corpus de este estudio de casos fueron producidas con una finalidad pedagógica, que se orienta a modelar de manera explícita la escritura de la demostración matemática a los estudiantes. Estas demostraciones corresponden a lo que Lew & Mejía-Ramos (2020) caracterizan como “una demostración producida en el contexto de un manual para estudiantes”, debido a su clara orientación a la enseñanza de la demostración como un género altamente especializado en la disciplina matemática. Estas demostraciones tienen una extensión y complejidad relativamente similar entre sí.

3.2 Procedimientos de análisis

En línea con los principios teóricos planteados por la LSF, un aspecto importante para la exploración de un significado es la consideración del contexto en el que funciona. En este sentido, el primer paso analítico de este estudio es la descripción lingüística del género. Para esta investigación, esto implica una descripción provisoria de la estructura esquemática de la demostración matemática. Para ello, se consideran fundamentalmente patrones de significado textual (PERIODICIDAD) e ideacional (FIGURA y CONEXIÓN). Los patrones de significado interpersonal se tienen en cuenta de manera selectiva y complementaria, en particular recursos de compromiso, que tienen alguna relevancia para los textos analizados.

Una vez identificada la estructura esquemática del género, el segundo paso analítico consiste en la identificación de secuencias en el estrato semántico-discursivo, lo que supone la determinación de figuras y conexiones que se establecen entre ellas. A la vez que se determinan las secuencias en el estrato semántico-discursivo, se identifican sus realizaciones léxico-gramaticales particulares. Este análisis permite revelar el potencial de diversificación gramatical de los significados semántico-discursivos de conexión en los datos analizados.

Dado que en las demostraciones se combina el lenguaje natural con el simbolismo matemático, es necesario clarificar cómo se entenderá en este artículo el simbolismo cuando se encuentra dentro de la estructura clausular del lenguaje natural. Expresiones de un símbolo, como ‘ x ’ o ‘ y ’, se consideran Participantes en el estrato léxico-gramatical y entidades en el

estrato semántico-discursivo, dado que concuerdan con el Proceso de la cláusula. Por su parte, una ecuación como ' $y=2m+1$ ' es una expresión que relaciona diferentes elementos mediante uno o más relatores (en este caso, $=$ y $+$). Así, puede entenderse como una figura de estado, dado que relaciona elementos. Sin embargo, cuando funciona dentro de la estructura clausal del lenguaje natural, la ecuación completa puede interpretarse como un Participante que es realizado por una expresión matemática que relaciona elementos, de modo similar a lo que ocurriría con una cláusula desplazada de rango. Esta interpretación se exemplifica en la Tabla 9.

Tabla 9—Lenguaje simbólico matemático y su interpretación lingüística

Participante ('grupo nominal')	Proceso	Participante ('cláusula desplazada de rango')
y	puede escribirse como	$y=2m+1$

Fuente: elaboración propia

Para considerar la diversificación grammatical (i.e. realizaciones congruentes e incongruentes), es necesario considerar tanto patrones semántico-discursivos como léxico-gramaticales. Para llevar a cabo esta distinción de manera sistemática se aplican al análisis los principios descriptivos de agnación y enación, originalmente propuestos por Gleason (1965). Dos realizaciones son consideradas agnados si involucran las mismas palabras y el mismo significado “nacional”, pero tienen distintas estructuras (e.g. ‘él la vio’; ‘ella fue vista por él’). Dos realizaciones son consideradas enados si tienen la misma estructura, pero involucran diferentes palabras (e.g. ‘él la vio’; ‘él la escuchó’). El principio de agnación –y, por consecuencia, el de enación– ha sido empleado para la identificación de realizaciones metafóricas (Hao, 2020; Heyvaert, 2003; Leiva, 2022). En este estudio, el principio de agnación se aplica predominantemente a la identificación de secuencias realizadas metafóricamente.

4 Resultados

Esta sección ofrece, en primer lugar, una descripción provisoria del género demostración matemática, a partir del conjunto de datos disponibles para este estudio (4.1), y luego una descripción de los patrones de CONEXIÓN identificados en ellos (4.2). A lo largo de la sección se utilizarán *íticas* para las fórmulas matemáticas y sus componentes, y VERSALITAS para señalar las conexiones.

4.1 Aproximación a la demostración matemática como un género

Desde la perspectiva de la LSF, se entiende que el propósito del género demostración matemática es proveer evidencia para un conocimiento matemático ya establecido, basándose en conocimiento previamente aceptado como verdadero y en procedimientos lógicos propios de la disciplina. Se propone que la estructura esquemática de este género se compone de tres etapas obligatorias:

Punto de la demostración \wedge Razonamiento matemático \wedge Confirmación

Esta estructura genérica se ejemplifica en la demostración matemática en la Tabla 10.

Tabla 10 – Estructura genérica de la demostración matemática

Etapa	Texto
Punto de la demostración	Supongamos que x e y son números impares. Demostremos que el producto xy es impar.
Razonamiento matemático	Como x es impar, sabemos que puede escribirse como $x = 2n + 1$, para algún entero n . En forma similar, que y sea impar implica que puede escribirse como $y = 2m + 1$, para algún entero m . Se sigue que $xy = (2n+1)(2m+1) = 4nm + 2n + 2m + 1 = 2(2nm + n + m) + 1$.
Confirmación	Concluimos que xy puede escribirse como $xy = 2k + 1$, para un cierto entero k , a saber, $k = 2nm + n + m$. Por lo tanto, xy es impar.

Fuente: elaboración propia

La primera etapa de la estructura esquemática de este género, **Punto de la demostración**, funciona para señalar la proposición matemática que será demostrada en el despliegue del texto. Para ello, se introducen los supuestos matemáticos que se toman como punto de partida y/o las condiciones que se considerarán para la demostración, y luego se presenta la proposición que debe demostrarse. En el ejemplo de la Tabla 9, el supuesto de entrada es la asunción de que 'x e y son números impares', mientras que la proposición a demostrar se reafirma en 'demostremos que el producto xy es impar'.

El ejemplo de la Tabla 9 presenta solamente una condición, lo que se relaciona con la orientación pedagógica de esta demostración y con el nivel de formación de los estudiantes a los que está dirigida. En otros casos, particularmente en demostraciones matemáticas solicitadas a estudiantes con una finalidad evaluativa, la etapa Punto de la demostración puede introducir más de una proposición que requiere ser confirmada. Cuando esto sucede, se produce una iteración secuencial de las etapas Razonamiento matemático y Confirmación. Un ejemplo de esta variación en el género se presenta en la Tabla 11. En este caso, la etapa Punto de la demostración es más compleja y se despliega a partir del enunciado de la tarea, expandiéndose hasta el inicio de la resolución.

Tabla 11 – Etapas iniciales de la demostración D2

Etapa	Texto
Punto de la demostración	Decimos que un número real es negativo si su inverso aditivo es positivo. Demuestre que el producto de un número positivo y uno negativo es un número negativo, y que el producto de dos números negativos es positivo.
Solución	Sean a un real positivo y b un real negativo cualesquiera. Por definición de número negativo, lo que se nos pide es demostrar que $-(a \cdot b)$ es positivo.
Razonamiento matemático 1	<p>Ahora bien,</p> $\begin{aligned} -(a \cdot b) &= (-1) \cdot (a \cdot b) && \text{(porque } (-1) \cdot x = -x \text{ para todo } x \text{)} \quad (1) \\ &= (a \cdot b) \cdot (-1) && \text{(comutatividad del producto)} \quad (2) \\ &= a \cdot (b \cdot (-1)) && \text{(asociatividad del producto)} \quad (3) \\ &= a \cdot (-b) && \text{(comutatividad y } (-1) \cdot x = -x \text{)} \quad (4) \end{aligned}$ <p>Por definición de número negativo, $-b$ es positivo. Esto implica que $a \cdot (-b)$ es positivo por ser producto de números positivos (axioma de la clausura de los números positivos con respecto al producto).</p>
Confirmación 1	Se demuestra así el resultado pedido
Razonamiento matemático 2	Si tanto a como b son negativos, entonces $(-a) \cdot (-b)$ es positivo por ser producto de números positivos. Por otro lado, $(-a) \cdot (-b) = ((-1) \cdot a) \cdot ((-1) \cdot b) = ((-1) \cdot (-1)) \cdot (a \cdot b) = 1 \cdot (ab) = ab$
Confirmación 2	Se concluye que ab es positivo, q.e.d ⁵ .

Fuente: elaboración propia

La segunda etapa del género, **Razonamiento matemático**, tiene como objetivo desarrollar el proceso lógico-matemático mediante el cual se puede llegar al punto de la demostración enunciado en la primera etapa. La etapa Razonamiento matemático se despliega mediante secuencias de figuras que integran el lenguaje natural con el simbolismo matemático. La Tabla 12 presenta las secuencias – es decir, la configuración ideacional de figuras y conexiones – que se despliegan en la etapa Razonamiento matemático en la demostración D1, introducida previamente.

Tabla 12 – Secuencias de figuras en etapa Razonamiento matemático en D1

conexiones	figuras
COMO	x es impar sabemos que puede escribirse como $[x = 2n + 1]$, para algún entero n .
EN FORMA SIMILAR,	que y sea impar
IMPLICA QUE	puede escribirse como $[y = 2m + 1]$ para algún entero m .
SE SIGUE QUE	$xy = (2n+1)(2m+1) = 4nm + 2n + 2m + 1 = 2(2nm + n + m) + 1$.

Fuente: elaboración propia

⁵ En la escritura de la demostración, esta sigla representa la cláusula ‘queda entonces demostrado’

En este ejemplo en particular, se observa una configuración de significados ideacionales interesante, que muestra la interacción entre lenguaje matemático y simbólico en la etapa Razonamiento matemático: la ecuación ‘ $x = 2n + 1$ ’ aparece como una entidad dentro de una figura de ocurrencia, instanciada en lenguaje natural (‘puede escribirse como $x=2n+1$ ’).

Una característica distintiva de los patrones ideacionales en la etapa Razonamiento matemático es la complejidad de las secuencias que se despliegan en ella, lo que va asociado al uso de recursos de CONEXIÓN para el establecimiento de la relación entre figuras. Un ejemplo de secuencias con mayor complejidad en la etapa Razonamiento matemático se da en la demostración D2, presentada en la Tabla 13.

Tabla 13 – Secuencias de figuras en etapa Razonamiento matemático (1 y 2) en D2

conexiones	figuras
Razonamiento matemático 1	
POR	definición de número negativo, $-b$ es positivo.
	Esto
IMPLICA	que $a \cdot (-b)$ es positivo
POR	ser producto de números positivos
Razonamiento matemático 2	
SI	tanto a como b son negativos
ENTONCES	$(-a) \cdot (-b)$ es positivo
POR	ser producto de números positivos
POR OTRO LADO,	$(-a) \cdot (-b) = ((-1) \cdot a) \cdot ((-1) \cdot b) = ((-1) \cdot (-1) \cdot (-1)) \cdot (a \cdot b) = 1 \cdot (ab) = ab$

Fuente: elaboración propia

Finalmente, la tercera etapa del género, **Confirmación**, funciona para establecer que la premisa inicial ha sido confirmada a través del razonamiento presentado. Esta etapa es la más breve en comparación con las dos anteriores. Desde un punto de vista textual, en esta etapa se recupera generalmente el punto de la demostración planteado en la etapa inicial del género. Un ejemplo de la etapa Confirmación se presenta en la Tabla 14.

Tabla 14 – Instancia de etapa Confirmación en D1

Etapa	Texto
Confirmación	CONCLUIMOS QUE xy puede escribirse como $xy = 2k + 1$, para un cierto entero k , a saber, $k=2nm+n+m$. POR LO TANTO, xy es impar

Fuente: elaboración propia

En este ejemplo, desde un punto de vista ideacional, la etapa Confirmación se instancia a partir de una figura de estado posicionada⁶, relacionada de manera explícita con la etapa anterior mediante la conexión ‘POR LO TANTO’. Esta etapa puede instanciarse también mediante otras configuraciones de significado ideacional, como a través de una figura de ocurrencia (‘Se demuestra así el resultado pedido’).

Otro ejemplo de etapa Confirmación se encuentra al final de la demostración D2 en la Tabla 15 a continuación:

Tabla 15–Instancia de etapa Confirmación en D2

Etapa	Texto
Confirmación	SE CONCLUYE QUE ab es positivo, q.e.d.

Fuente: elaboración propia

La descripción inicial del género demostración matemática aquí presentada permite aproximarse al estudio de la CONEXIÓN ‘por arriba’, teniendo en cuenta las motivaciones que provienen del estrato del género. Tal como se sugiere en esta descripción, las etapas Razonamiento matemático y Confirmación serían las dos más productivas para explorar la CONEXIÓN, teniendo en cuenta su función en la estructura esquemática del género.

4.2 La conexión en la demostración matemática

En esta sección, se presentan los resultados del análisis de los recursos de CONEXIÓN que se despliegan en las demostraciones matemáticas analizadas. Se considera aquí, en primer lugar, el tipo de conexión; luego, las opciones de realización congruente e incongruente, y, finalmente, las opciones en los planos externo e interno que se identifican en los datos. Se expondrán dos instancias como ejemplos para cada caso y luego se desplegará el análisis detallado de una de ellas, para profundizar en la configuración.

4.2.1 Tipos de significado en la conexión

En las demostraciones matemáticas analizadas aquí, el tipo de conexión más recurrente es el consecuencial. Este patrón es consistente con lo planteado tanto en la literatura matemática como en aproximaciones desde la didáctica de la matemática y la lingüística a los textos escritos en esta área (por ejemplo, Alfaro-Carvajal *et al*, 2019; Lew & Mejía-Ramos, 2020; O’Halloran, 2005)

⁶ Desde un punto de vista ideacional, y según la propuesta de Hao (2020), las figuras pueden ser expandidas mediante varios recursos. Uno de ellos es la posición de figura, a partir de la cual los procesos verbales y mentales que proyectan una cláusula se interpretan en el estrato semántico-discursivo como recursos para posicionar una figura. En este sentido, no constituyen una figura en sí mismos, sino que la expanden. Desde una mirada interpersonal, la posición se asocia con el compromiso de heteroglosia (Martin & White, 2005), ya que reconoce la presencia de otras voces en el discurso.

Respecto de las conexiones de tipo consecuencial, se distinguen en los datos dos sub-tipos: conexión consecuencial causal y conexión consecuencial de condición. La primera establece una relación lógica en la que se introduce una condición necesaria para que un cierto efecto ocurra. El segundo tipo de conexión consecuencial señalado establece una relación entre un efecto y una condición bajo la cual este puede suceder (Martin, 1992; Martin & Rose, 2007). Las secuencias (1) y (2) presentan ejemplos de conexión consecuencial causal (en adelante, simplemente 'conexión causal') y conexión consecuencial de condición, respectivamente.

- (1) COMO x es impar, sabemos que puede escribirse como $x = 2n + 1$ (D1)
- (2) si tanto a como b son negativos, ENTONCES $(-a) \cdot (-b)$ es positivo (D2)

Mirado desde el estrato semántico-discursivo, el ejemplo (1) introduce una secuencia en la que se despliegan dos figuras (señaladas con números romanos), que se relacionan entre sí mediante la conexión 'COMO':

COMO	(i)	x es impar
	(ii)	(sabemos que) puede escribirse como $x = 2n + 1$.

Por su significado, la conexión 'COMO' entre estas figuras establece a una de ellas como causa (x es impar) y a la otra como un efecto ('puede escribirse como $x = 2n + 1$ '). La segunda figura de la secuencia presenta además una posición ('sabemos que'), que funciona para introducir la fuente de 'lo sabido' (Hao, 2020).

En el caso del ejemplo (2), la secuencia de figuras se constituye por medio de dos conectores:

SI	(i)	tanto a como b son negativos
ENTONCES	(ii)	$(-a) \cdot (-b)$ es positivo

La relación que las conexiones 'SI' y 'ENTONCES' establecen entre estas figuras posiciona a la figura (i) como causa y la figura (ii) como efecto.

Tanto el ejemplo (1) como (2) corresponden a realizaciones congruentes de la conexión. La realización congruente de este tipo de significado se profundiza en el siguiente apartado

4.2.2 Congruencia de la conexión

En los datos analizados, las conexiones causales en una secuencia pueden realizarse tanto de manera congruente como incongruente. La realización **congruente** de la conexión se ilustra en los ejemplos (1) y (2) presentados más arriba. En estos casos, las figuras que componen las secuencias son realizadas 'por abajo' por cláusulas (o complejos de cláusulas) y la conexión se realiza mediante un conector, tal como se muestra a continuación. Esta representación alternativa del ejemplo (1) busca identificar las unidades semántico-discursivas que constituyen esta secuencia y las unidades léxico-gramaticales que las realizan:

Tabla 16 – Realización congruente de secuencia: complejo clausular

semántica-discursiva	secuencia		
	conexión causal	figura causa	figura efecto posicionada
	COMO	x es impar	sabemos que puede escribirse como $x = 2n + 1$
léxico-gramática	conector	cláusula	complejo clausular
		complejo clausular	

Fuente: elaboración propia

Teniendo en cuenta la perspectiva trinocular propuesta, esta secuencia puede abordarse ‘por alrededor’, desde la metafunción textual. Desde este punto de vista, en esta secuencia, la figura causa es realizada por una cláusula hipotáctica que precede a la cláusula independiente, por lo que se considera Tema marcado, de manera similar a las Circunstancias (Moyano, 2021, p. 497; Martin, Quiroz & Wang, 2023, p. 344). El análisis de la distribución de Tema y Nuevo en las cláusulas que realizan la secuencia se presenta en la Tabla 17.

Tabla 17 – Análisis de Tema en cláusulas realizando secuencia 1

Tema textual	Tema tópico marcado	Tema tópico no marcado	Proceso	Nuevo
	COMO x es impar	(nosotros)	sabemos	
que		(x)	puede escribirse	como $x = 2n + 1$

Fuente: elaboración propia

El análisis temático muestra cómo la figura causa, realizada por una cláusula hipotáctica en posición inicial, funciona textualmente como un Tema marcado, dejando como Tema de la cláusula principal a un participante concordante ‘(nosotros)’ que aparece de manera implícita y es recuperado por la flexión verbal del proceso de la cláusula mental. Esta cláusula mental proyecta una segunda cláusula, cuyo Tema es ‘(x)’, también recuperado textualmente, que corresponde a la variable matemática relevante en esta secuencia. El hecho de que la figura causa aparezca como Tema tópico marcado permite, en este caso, posicionarla como información que se asume conocida, debido a que ha sido presentada previamente, y dar relevancia informativa al Nuevo, que es la forma “nueva” de escribir x como $2n + 1$. Desde el punto de vista de la estructura esquemática del género, el posicionamiento del Tema tópico marcado ‘Como x es impar’ permite identificar el cambio de etapa: del Punto de la demostración al Razonamiento matemático

Otra manera de organizar la información de la secuencia (1) no tendría los mismos efectos textuales señalados. Por ejemplo, puede proponerse una forma agnada de realizar este significado ideacional, en la que se altera el orden de las cláusulas que realizan las figuras, de modo que la figura causa aparece pospuesta a la figura efecto, usando un conector:

- YA QUE (i) Sabemos que x puede escribirse como $x=2n+1$

Para mantener el tipo de conexión causal en el agnado, se emplea un conector cuyo significado permite introducir la figura causa (en vez de la figura efecto) como si fuera información nueva. En este caso, se opta por el conector 'ya que', pero otras alternativas podrían ser 'pues', 'puesto que', 'porque', etc. En este sentido, la propuesta del agnado supone mantener el tipo de significado que vincula dos figuras en una secuencia, para lo cual puede seleccionarse cualquier opción de conector adecuada para el significado en juego. El análisis de Tema y Nuevo en el agnado se presenta en la Tabla 17.1.

Tabla 17.1 – Análisis de Tema en cláusulas realizando agnado (1.1)

Tema textual	Tema tópico marcado	Tema tópico	Proceso	Nuevo
		(nosotros)	sabemos que	
YA QUE		(x)	puede escribirse	como $x = 2n+1$
		x	es	impar

Fuente: elaboración propia

Desde la mirada ‘por alrededor’, cada una de estas secuencias tiene una distribución de la información diferentes, lo que puede tener consecuencias a nivel del texto y, por lo tanto, de la lógica de la demostración. En el caso de la secuencia agnada (1.1), la cláusula *x es impar* es introducida como si fuera información nueva, lo que reorganiza la manera en que se despliega la información en este texto. En este sentido, las opciones de realización de la CONEXIÓN tienen un impacto en la forma en la que los pulsos de información van construyendo a lo largo del texto la lógica de la demostración.

Además de realizaciones congruentes, los datos analizados permiten identificar dos patrones de realización **incongruente o metafórica** de la conexión, cada uno asociado a un recurso lingüístico particular: por una parte, el uso de preposición, y por otra, el uso de grupo verbal.

4.2.2.1 Conexión incongruente a través de preposición

Un primer recurso para la realización incongruente de la conexión corresponde a la preposición 'por'. En estos casos, la secuencia es realizada por una sola cláusula, en la que pueden observarse diferentes configuraciones estructurales. Algunos ejemplos de esta realización incongruente de la conexión por preposición se presentan a continuación en (3) y (4).

- (3) (i) $a \cdot (-b)$ es positivo
 POR (ii) ser producto de números positivos

(4) POR (i) definición de número negativo,
 (ii) $-b$ es positivo

En (3) se observa una secuencia compuesta por dos figuras de estado. La secuencia es realizada, en el estrato léxico-gramatical, por una cláusula que incluye el Proceso relacional ‘es’, como se muestra en la a continuación.

Tabla 18 – Estructura clausular de secuencia realizada de manera incongruente

léxico-gramática	a · (–b)	es	positivo	POR [[ser producto de números positivos]]
	Participante	Proceso	Participante	Circunstancia
cláusula				

Fuente: elaboración propia

Desde el punto de vista léxico-gramatical, la Circunstancia de la cláusula está constituida por la preposición ‘por’ y la cláusula incrustada ‘[[ser producto de números positivos]]’. Mirado ‘por arriba’, desde el estrato semántico-discursivo, la Circunstancia realiza una figura, que establece una conexión con la figura anterior mediante la preposición ‘por’. Esta preposición funciona como conexión causal, estableciendo a la primera figura como una consecuencia y a la segunda como su causa. Esta relación lógica entre las dos figuras se presenta en la Tabla 19.

Tabla 19 – Realización metafórica de secuencia: cláusula

semántica-discursiva	secuencia				
	figura efecto			conexión causal	figura causa
	a · (–b)	es	positivo	POR	ser producto de números positivos
léxico-gramática	participante	proceso	participante	circunstancia	
	cláusula				

Fuente: elaboración propia

La Tabla 19 permite apreciar la tensión inter-estratal que se produce entre las unidades léxico-gramaticales y semántico-discursivas en (3): una secuencia, que congruentemente se realizaría en un complejo clausular se realiza en este caso dentro de una cláusula. La Tabla 1.1 presenta el agnado congruente de la secuencia (3). En este agnado ambas figuras son realizadas por cláusulas en rango:

Tabla 19.1 – Realización congruente de secuencia: complejo clausular (forma agnada)

semántica-discursiva	secuencia		
	figura efecto	conexión causal	figura causa
	$a \cdot (-b)$ es positivo	YA QUE	es producto de números positivos
léxico-gramática	cláusula	conector	cláusula
	complejo clausular		

Fuente: elaboración propia

En la forma agnada mostrada en la Tabla 19.1, la cláusula incrustada que funcionaba como parte de la Circunstancia en (3) ('ser producto de números positivos') es realizada como la cláusula en rango 'es producto de números positivos' y la conexión pasa de realizarse a través de la preposición 'POR' a ser realizada por el conector 'YA QUE'.

4.2.2.2 Conexión incongruente a través de grupo verbal

En las demostraciones matemáticas analizadas, otro recurso relevante para la realización incongruente o metafórica de la conexión corresponde al grupo verbal. En los ejemplos (5) y (6), se presentan dos casos en los que la conexión de las secuencias se realiza mediante el Proceso 'implica':

- (5) (i) que y sea impar
 IMPLICA (ii) que puede escribirse como $y=2m+1$, para algún entero m
- (6) (i) Esto
 IMPLICA (ii) que $a \cdot (-b)$ es positivo

Si observamos en detalle la secuencia (5), podemos ver que las dos figuras conectadas por 'implica' son encapsuladas en el rango del grupo a través de cláusulas incrustadas: 'que y sea impar' y 'que puede escribirse como $y=2m+1$ '. La conexión es realizada a través del grupo verbal 'IMPLICA', descrito como causal en otros discursos, como el de la historia (Leiva, 2022). En este caso, al igual que el descrito en (3), la secuencia semántico-discursiva es realizada léxico-gramaticalmente por medio de una sola cláusula, lo que supone una tensión inter-estatal. Esta realización incongruente de la secuencia se representa a continuación.

Tabla 20–. Realización metafórica de secuencia: cláusula

secuencia		
figura causa	conexión causal	figura efecto
[[que y sea impar]]	IMPLICA	[[que puede escribirse como $y = 2m + 1$, para algún entero m]]
cláusula		
Participante grupo nominal	Proceso grupo verbal	Participante grupo nominal

Fuente: elaboración propia

La secuencia (5) puede desempaquetarse como un agnado congruente. En esta realización alternativa ambas figuras son realizadas no como cláusulas incrustadas sino como cláusulas en rango, y la conexión no como un grupo verbal sino como un conector, en este caso, 'entonces'. El agnado congruente de la secuencia (5) se muestra a continuación, en la Tabla 20.1.

Tabla 20.1. Realización congruente de secuencia: complejo clausular (agnado)

secuencia		
figura causa	conexión causal	figura efecto
y es impar	ENTONCES	puede escribirse como $y = 2m + 1$, para algún entero m
complejo clausular		

Fuente: elaboración propia

Al observar en detalle la secuencia (6) antes introducida, vemos que solo una de las figuras conectadas por 'IMPLICA' es encapsulada en el rango del grupo a través de una cláusula incrustada ('que $a \cdot (-b)$ es positivo'), mientras que la otra es encapsulada a través de una referencia textual anafórica ('Esto').

La secuencia (6) también puede desempaquetarse como un complejo clausular. La realización alternativa hace la secuencia explícita, al realizar ambas figuras como cláusulas (explicitando la figura referida por el ítem textual 'Esto') y la conexión como la conjunción 'ENTONCES. La secuencia metafórica y su agnado congruente se muestran en la Tabla 21 y la Tabla 21.1 respectivamente a continuación:

Tabla 21 – Realización metafórica de secuencia: cláusula

secuencia		
figura causa	conexión causal	figura efecto
Esto	IMPLICA	[[que $a \cdot (-b)$ es positivo]]
Participante grupo nominal	Proceso grupo verbal	Participante grupo nominal
cláusula		

Fuente: elaboración propia

Tabla 21.1 – Realización congruente de secuencia: complejo clausular (agnado)

secuencia		
figura causa	conexión causal	figura efecto
$-b$ es positivo	ENTONCES	$a \cdot (-b)$ es positivo
cláusula	conector	cláusula
complejo clausular		

Fuente: elaboración propia

Preliminarmente, podemos plantear algunas diferencias entre la manera en que se despliega la conexión incongruente en las demostraciones matemáticas analizadas y las descripciones disponibles para otras disciplinas, en particular para el discurso de la historia. En las demostraciones analizadas, el empaquetamiento de las figuras en el rango del grupo no ocurre por medio de nominalización, sino por medio de una cláusula incrustada, como en (5), o mediante referencia textual, como en (6). Esto tiene como una consecuencia relevante que no hay en la demostración un potencial evaluativo como el que se abre con la nominalización (e.g. **el importante estatus de impar implica su inmediata escritura como...*). Esto está en línea con lo que O'Halloran (2015) y Doran (2018b) describen con respecto a la ausencia de componente interpersonal en el discurso matemático. En efecto, a nivel interpersonal, no se identifican en las demostraciones analizadas el uso de Procesos que gradúen la fuerza causal, como 'estimular', 'contribuir' o 'permitir' (cf. Leiva, 2022; Leiva & Oteíza, 2023).

A nivel textual, sin embargo, el uso de la metáfora lógica tal y como se despliega en el discurso matemático analizado parece traer ganancias. En este caso, la encapsulación de la primera figura en el rango del grupo posibilita la tematización de una figura previamente introducida, de una manera similar a la cláusula hipotáctica antepuesta en la secuencia (1). A continuación, se representa el análisis de la organización de la información en la secuencia (5) en contraste con su agnado congruente, en las Tablas 22 y 22.1 respectivamente:

Tabla 22 – Análisis de Tema y Nuevo de secuencia (metafórica)

Tema	Proceso	Nuevo
[[que y sea impar]]	IMPLICA	[[que puede escribirse como $y = 2m + 1$, para algún entero m]]

Fuente: elaboración propia

Tabla 22.1 – Análisis de Tema y Nuevo de la secuencia agnada (congruente)

Tema textual	Tema	Proceso	Nuevo
	y	es	impar
ENTONCES	(y)	puede escribirse	como $y = 2m + 1$, para algún entero m

Fuente: elaboración propia

Tal como se observa en la secuencia y su agnado, las opciones de realización incongruente de la conexión permiten en el flujo textual posicionar cierta información como un punto de partida (Tema) y otra como información relevante (Nuevo). Esto tiene consecuencias respecto de la manera en que se despliega la lógica argumentativa propia del género demostración matemática.

4.2.3 Conexión externa versus conexión interna

En las demostraciones matemáticas analizadas es posible identificar conexiones que funcionan tanto en el plano externo como en el plano interno. En los ejemplos (1-4) ya introducidos, las conexiones corresponderían a conexiones externas, que existen ‘fuera del texto’, como conocimiento matemático aceptado como verdadero en el campo:

- (1) COMO (i) x es impar
(ii) sabemos que puede escribirse como $x = 2n + 1$
- (2) SI (i) tanto a como b son negativos,
ENTONCES (ii) $(-a) \cdot (-b)$ es positivo
- (3) (i) $a \cdot (-b)$ es positivo
POR (ii) ser producto de números positivos
- (4) POR (i) definición de número negativo,
(ii) $-b$ es positivo

En estas secuencias, la conexión vincula solamente dos figuras para formar una secuencia. Dado que la conexión en todos estos casos es de tipo causal, una de las figuras se identifica como causa (‘ser producto de números positivos’) y la otra figura se identifica como efecto (‘ $a \cdot (-b)$ es positivo’). Esta conexión causal entre las proposiciones se entiende

como experiencial, en tanto existe “fuera del texto”, al ser parte del conocimiento matemático. El hecho de que se trate de conocimiento matemático del campo permite agregar a la figura efecto en las secuencias (2-4) arriba una posición (‘sabemos que’), que contribuye a señalar de manera explícita a esta figura como una premisa sabida y compartida a partir del conocimiento del campo, tal como aparece en la secuencia (1). La posición permite de esta forma corroborar la naturaleza externa de la conexión en juego en cada secuencia, tal como se muestra a continuación:

figura causa	conexión causal	figura efecto posicionada
SI tanto a como b son negativos	ENTONCES	sabemos que $(-a) \cdot (-b)$ es positivo
figura efecto posicionada	conexión causal	figura causa
Sabemos que $a \cdot (-b)$ es positivo	POR	ser producto de números positivos.
conexión causal	figura causa	figura efecto posicionada
POR	definición de número negativo	sabemos que $-b$ es positivo

Por otro lado, también es posible identificar en el discurso matemático conexiones de naturaleza interna, tal y como se ha destacado en la literatura previa (O’Halloran, 2005). Estas conexiones funcionan retóricamente para organizar el razonamiento y el despliegue de la información en el texto. Un primer patrón relevante de conexión interna en las demostraciones analizadas es que esta conecta fragmentos más extensos de texto que la conexión externa, que pueden corresponder a dos secuencias o a una figura y una secuencia entre sí. En (7), de la demostración D1, se muestra un ejemplo de conexión interna entre dos secuencias, mediante el conector ‘EN FORMA SIMILAR’:

(7) COMO	(i) x es impar	secuencia (I)
	(ii) sabemos que puede escribirse como $x = 2n + 1$, para algún entero n .	
EN FORMA SIMILAR,		
	(iii) que y sea impar	secuencia (II)
IMPLICA	(iv) que puede escribirse como $y = 2m + 1$, para algún entero m	

El ejemplo (7) presenta una comparación entre dos secuencias causales externas, que son parte de las proposiciones conocidas por conocimiento matemático –correspondientes a las secuencias (1) y (5) previamente introducidas. La similitud es aquí retóricamente establecida entre dos razonamientos que se presentan como ejemplos de un mismo principio matemático. En este caso, el conector que relaciona las dos secuencias es de significado comparativo de similitud, según la categorización propuesta por Martin (1992). Una particularidad de las conexiones comparativas internas es que permiten la intercambiabilidad del orden de las unidades que vinculan, es decir, una reorganización del flujo de la información desde el punto de vista textual. De esta manera, es posible modificar la forma en que se organiza la

información en este complejo, intercambiando el orden de las secuencias. Esto permite construir un significado agnado de la secuencia original. Estas secuencias, la original y su agnado, se muestran en las tablas a continuación:

Tabla 23 – Complejo de secuencias con conexión interna

secuencia con conexión causal	conexión interna comparativa	secuencia con conexión causal
Como x es impar, sabemos que puede escribirse como $x = 2n + 1$, para algún entero n .	EN FORMA SIMILAR	que y sea impar IMPLICA que puede escribirse como $y = 2m + 1$, para algún entero m

Fuente: elaboración propia

Tabla 23.1 – Complejo de secuencias con conexión interna: reorganización textual

secuencia con conexión causal	conexión interna comparativa	secuencia con conexión causal
Que y sea impar IMPLICA que puede escribirse como $y = 2m + 1$, para algún entero m	EN FORMA SIMILAR	Como x es impar, sabemos que puede escribirse como $x = 2n + 1$, para algún entero n .

Fuente: elaboración propia

En los ejemplos señalados, la conexión interna funciona para relacionar secuencias simples, en las que se relacionan solamente dos figuras por una relación de conexión externa. Sin embargo, a medida que aumenta la complejidad ideacional de la etapa Razonamiento matemático, es posible observar en estas secuencias un anidamiento de conexión, es decir, la conformación de secuencias dentro de secuencias. Este patrón puede ser interpretado en términos de complejos de secuencias. En los complejos de secuencias es posible identificar dos o más relaciones de conexión externa, realizada explícita o implícitamente. La realización implícita de la conexión ha sido señalada como un rasgo del discurso matemático por otros autores en el ámbito de la LSF (O'Halloran, 2005, p. 119; Schleppegrell, 2007, p. 139).

Las secuencias que constituyen un complejo con otra secuencia pueden realizarse tanto mediante lenguaje natural como a través de lenguaje matemático. Un ejemplo de este patrón se presenta en la secuencia (8) a continuación, que conecta internamente los complejos de secuencia (I) y (II).

(8) SI	(i) tanto a como b son negativos	complejo de secuencia (I)	
ENTONCES	(ii) $(-a) \cdot (-b)$ es positivo		
POR	(iii) ser producto de números positivos		
POR OTRO LADO			
(i) $(-a) \cdot (-b) = ((-1) \cdot a) \cdot ((-1) \cdot b)$			
(ENTONCES) (ii) $= ((-1) \cdot (-1) \cdot (-1)) \cdot (a \cdot b)$			
(ENTONCES) (iii) $= 1 \cdot (ab)$			
(ENTONCES) (iv) $= ab$			

En el ejemplo (8), el complejo de secuencias (I) presenta un conjunto de premisas matemáticas que explican de manera condensada a través de lenguaje natural el cambio de polaridad en el producto de dos números negativos. En este complejo se observa un ejemplo de anidación de secuencias: las figuras (ii) y (iii) forman una secuencia causal mediante la conexión ‘POR’, y esa secuencia se conecta con la figura (i) en una relación consecuencial de condición mediante las conexiones ‘SI... ENTONCES’ para formar el complejo de secuencia (I). Estas relaciones de anidamiento se representan en 8 con el uso de diferentes sangrías. La conexión interna ‘POR OTRO LADO’⁷ introduce el complejo de secuencia (II), que despliega el principio matemático presentado en el complejo (I) a través de lenguaje simbólico, para demostrar de forma explícita y paso por paso por qué el producto de dos números negativos es positivo. En el complejo de secuencia (II), cada una de las ecuaciones o igualdades que se presenta es interpretada lingüísticamente como una figura de estado, ya que, desde el punto de vista del lenguaje natural, pueden ser leídas como cláusulas relationales del tipo ‘menos a por menos b es igual a menos 1 por a , por menos uno por b ’ (O’Halloran, 2005). La disposición de las ecuaciones en una línea hacia abajo se entiende aquí como conexión causal implícita expresada mediante un conector causal entre paréntesis: (ENTONCES).

En los ejemplos presentados hasta este punto, la conexión interna se realiza mediante conectores. Sin embargo, al igual que las conexiones externas, las conexiones internas también pueden ser realizadas por grupos verbales en las demostraciones matemáticas estudiadas. Un ejemplo de esto se presenta en la secuencia (9), a continuación.

(9)	COMO	(i) x es impar	secuencia (I)
	ENTONCES	(ii) sabemos que puede escribirse como $x = 2n + 1$, para algún entero n .	
EN FORMA SIMILAR			
		(i) que y sea impar	secuencia (II)
	IMPLICA	(ii) que puede escribirse como $y = 2m + 1$, para algún entero m	
	SE SIGUE QUE	(iii) $xy = (2n+1)(2m+1) = 4nm + 2n + 2m + 1 = 2(2nm + n + m) + 1$.	secuencia (III)
	CONCLUIMOS QUE	(iv) xy puede escribirse como $xy = 2k + 1$, para un cierto entero k , a saber, $k = 2nm + n + m$.	secuencia (IV)

⁷ Si bien el significado de ‘POR OTRO LADO’ corresponde regularmente a un significado comparativo de similitud (Martin & Rose, 2007), en este caso la relación de significado que se establece es de carácter más bien elaborativo, tal como queda demostrado por el complejo de secuencias II.

El ejemplo (9) presenta algunas de las opciones que se identifican en las demostraciones para realizar la conexión interna mediante grupo verbal. En este caso, aparecen los procesos 'SE SIGUE' y 'CONCLUIMOS' para establecer este tipo de relaciones. Estas conexiones funcionan para organizar el argumento, mostrando las deducciones lógicas establecidas retóricamente por el escritor de la demostración. Su contribución retórica puede verse en relación con el género: estas conexiones internas indican el cierre de una etapa y el inicio de otra: 'SE SIGUE' permite concluir la etapa de Razonamiento matemático y 'CONCLUIMOS que' señala la etapa Confirmación, que completa la estructura esquemática del género. Nuevamente, las unidades vinculadas por estas conexiones internas son mayores a las unidades vinculadas típicamente por las conexiones externas (i.e. cláusulas); en este caso, los límites de la unidad que cumple la función efecto no es identificable de manera categórica, es decir, no se puede señalar con certeza su alcance exacto en el flujo del texto. Otro rasgo de las conexiones causales internas identificadas en las demostraciones es que son realizadas por Procesos de tipo mental y pueden ser atribuibles al hablante que produce el texto, lo que hace más evidente el rol retórico que cumplen.

El análisis de las demostraciones matemáticas exploradas ha permitido identificar las diferentes opciones de CONEXIÓN que se despliegan en este género, así como la manera en que este significado lógico se relaciona con la estructura esquemática del género que se instancia en la demostración. La perspectiva trinocular asumida ha permitido describir estos patrones de CONEXIÓN teniendo en cuenta su interacción con opciones textuales ('por alrededor'), sus diferentes realizaciones léxico-gramaticales ('por abajo') y su despliegue en relación con la estructura esquemática del género ('por arriba'). Las características de las opciones de CONEXIÓN a partir de esta mirada trinocular se sintetizan en la Tabla 24.

Tabla 24 – Patrones de CONEXIÓN desde una mirada trinocular: síntesis

Tipo de conexión	Plano de conexión	Unidades semántico-discursiva conectadas	Realización léxico-gramatical ('por abajo')	Potencial de organización textual ('por alrededor')	Vínculo con género ('por arriba')
Causal	Externa	Figuras	Congruente: mediante conectores Incongruente: mediante grupo verbal y preposición	Organización textual modificable con cambio de conector para mantener sentido de relación causa-efecto	Establece relaciones en las etapas Punto de la demostración y Razonamiento matemático
Causal	Interna	Secuencias o complejos de secuencias	Congruente: mediante conectores y grupo verbal	Organización textual no modificable	Señala límites de las etapas Razonamiento matemático y Confirmación
Comparativa	Interna	Secuencias	Congruente: mediante conectores	Organización textual modificable sin cambio de conector	Elabora relaciones para el despliegue de la etapa Razonamiento matemático

Fuente: elaboración propia

5 Conclusiones

El objetivo de este trabajo fue ofrecer una descripción inicial de la demostración como género discursivo y mostrar los patrones de CONEXIÓN que son relevantes en este, desde la perspectiva teórica de la LSF. Este estudio propone una exploración a un campo disciplinar y a un género específico que desde la investigación lingüística en español ha sido en general escasamente explorado. Basándose en las descripciones sistémico-funcionales previas del discurso matemático, este estudio busca comprender cómo funciona el lenguaje natural que acompaña al simbolismo matemático en uno de los géneros más importantes en la formación de matemáticas y matemáticos como parte de su progresiva especialización en la disciplina. Esta descripción se ha propuesto desde una perspectiva teórica y metodológica de naturaleza funcional, que permite abordar el estudio de la CONEXIÓN en relación con otras opciones de significado. En el marco de la teoría sistémica, esta aproximación se ha entendido como una mirada trinocular, que involucra observar los significados desde diferentes estratos y sistemas complementarios.

En consideración de la perspectiva trinocular, el punto de partida para esta exploración ha sido la descripción de la demostración matemática como un género discursivo. Esta descripción ha considerado los recursos textuales e ideacionales que permiten identificar patrones o síndromes de significado asociados a las diferentes etapas que se despliegan en el texto para cumplir su propósito social. Los datos explorados permiten proponer una descripción inicial de la estructura esquemática de tres etapas: Punto de la demostración \wedge Razonamiento matemático \wedge Confirmación. La definición de esta estructura esquemática complementa, desde una mirada profunda a los patrones de lenguaje, las descripciones de este género propuestas desde el ámbito de la enseñanza de las ciencias. El aporte de este estudio radica, en este sentido, en ofrecer criterios lingüísticos específicos y explícitos para comprender cómo se organiza una demostración matemática desde el punto de vista semiótico.

El interés por una exploración lingüística de la CONEXIÓN en este género ha sido motivado por la gran relevancia que se otorga en la literatura a las relaciones lógicas en la demostración (Alvarado & González, 2009; Camacho *et al.*, 2014; O'Halloran, 2005; Schleppegrell, 2007; Spanos *et al.*, 1988). El análisis lingüístico confirma que las conexiones constituyen un recurso crucial para dar sentido a la relación entre proposiciones y expresiones matemáticas, así como para explicitar el proceso deductivo en todas las etapas de la demostración. El análisis ha permitido identificar conexiones causales y comparativas, realizaciones congruentes e incongruentes, así como conexiones en el plano externo e interno; específicamente, conexiones causales externas, causales internas y comparativas internas, con formas de realización variadas. Estas incluyen conectores o frases conectivas, preposiciones y grupos verbales. Los patrones identificados en este estudio son consistentes con algunos de las características propuestas por O'Halloran (2005) para el significado lógico en el discurso matemático, como la presencia de relaciones lógicas de tipo condicional y causal, y de relaciones cohesivas, es decir, conexiones internas. Un patrón interesante en los datos es que la conexión interna es siempre congruente, incluso cuando se realiza mediante grupo verbal (e.g. 'SE SIGUE QUE'), ya que no condensa la realización de una secuencia al rango de la cláusula. Las opciones en el plano externo, en cambio, sí son en ocasiones realizadas de

manera incongruente (e.g. 'POR' e 'IMPLICA' en ejemplos 3-6), es decir, pueden funcionar para condensar una secuencia en el rango de la cláusula.

Desde una mirada 'por alrededor', la interacción de las opciones de CONEXIÓN con recursos de la metafunción textual permite identificar algunas implicancias importantes de optar por ciertos recursos por sobre otros. Respecto de las opciones de conexión causal congruente, se observa que, en el plano externo, un cambio en el flujo informativo afecta el orden de la relación causa-efecto entre las figuras y puede tener implicaciones respecto de qué información se presupone como punto de partida. En el caso de la conexión causal en el plano interno, se observa la imposibilidad de reorganizar el flujo textual sin que se afecte la lógica de su despliegue, mientras que en el caso de la conexión comparativa en el plano interno es posible intercambiar el orden de las figuras sin alterar el sentido de la secuencia.

Otro aspecto interesante desde la mirada 'por alrededor' involucra los significados interpersonales en el estrato semántico-discursivo. Tal como señalan algunos autores en el ámbito de la LSF, el componente interpersonal pareciera ser irrelevante en el discurso matemático (Doran, 2018b; O'Halloran, 2005). Esto se observa en patrones de CONEXIÓN en la demostración, en particular respecto de la realización incongruente de la conexión externa. En los datos analizados, el uso de grupos verbales como recurso para la CONEXIÓN no cumple un rol en la gradación de la fuerza causal (e.g. 'CAUSAR' vs. 'PERMITIR'), como sí ha sido descrito en el discurso de otras disciplinas, como la historia (Coffin, 2004; Leiva & Oteíza, 2023). Sin embargo, en los datos se observan otros patrones interpersonales, como el uso de posiciones (e.g. 'sabemos que'), que parecen sugerir algún rol de los significados interpersonales. Este es un aspecto que requiere mayor exploración desde una perspectiva lingüística.

En este artículo, el estudio de la demostración matemática se ha focalizado en demostraciones elaboradas en el contexto de educación terciaria, cuestión que amplía la naturaleza de los datos estudiados hasta ahora en las investigaciones en el marco sistémico-funcional. Una característica relevante de estas demostraciones es que han sido elaboradas por docentes universitarios de matemáticas con una finalidad pedagógica. Se trata de textos que buscan presentarse como buenos modelos de la manera en que se espera que los estudiantes elaboren el razonamiento matemático y desplieguen los recursos semióticos para construir la demostración como un texto cohesivo. De esta manera, los datos estudiados aquí pueden considerarse como modelos prototípicos de demostración, para un cierto nivel de formación y un cierto nivel de complejidad disciplinar. En este sentido, la exploración de la demostración matemática que aquí se presenta tiene un gran potencial de expansión, para considerar diferentes momentos formativos de los estudiantes y diferentes contextos en los que los especialistas producen demostraciones para crear conocimiento nuevo.

Este trabajo, basado en un corpus acotado de demostraciones, busca ser un punto de partida para desarrollar descripciones teóricamente fundadas de géneros del ámbito de las matemáticas. Un aspecto clave que ha permitido llevar a cabo esta exploración lingüística ha sido el trabajo interdisciplinario con docentes de matemáticas, no solo para acceder a los datos analizados sino sobre todo para entender, desde una mirada especializada en el campo de las matemáticas, qué es la demostración y cómo se espera que funcione en el contexto de la disciplina. Esta colaboración se ha orientado, especialmente, a elaborar una mirada lingüística sobre la demostración que pueda tener aplicaciones pedagógicas para su enseñanza, teniendo como fundamento el principio de aplicabilidad propuesto por la lingüística sistémico funcional.

Agradecimientos

A los docentes de la Facultad de Matemáticas de la Pontificia Universidad Católica Mahsa Allahbakhshi, Ricardo Menares, Duvan Henao y Antonio Behn, por el fructífero trabajo en torno a las demostraciones matemáticas que sostuvieron con una de las autoras. Esta colaboración interdisciplinaria fue, en definitiva, la motivación original de este estudio.

Declaración de autoría

Declaramos que la escritura del artículo fue realizada por las dos autoras en conjunto. En términos del análisis, el de género fue liderado por la autora Margarita Vidal Lizama, mientras que el de CONEXIÓN fue liderado por la autora Natalia Leiva Salum.

Referencias

- ACURSO, K.; GEBHARD, M.; PURINGTON, S.B. Analyzing Diverse Learners' Writing in Mathematics: Systemic Functional Linguistics in Secondary Pre-Service Teacher Education. *International Journal for Mathematics Teaching and Learning*, v. 18, n. 1, p. 84-108, 2017. DOI: <https://doi.org/10.4256/ijmtl.v18i1.48>
- ADAMS, T. L. Reading mathematics: More than words can say. *The Reading Teacher*, v. 56, n.8, p. 786–795, 2003. DOI: <http://www.jstor.org/stable/20205297>.
- ALFARO-CARVAJAL, C., FLORES-MARTINEZ, P.; VALVERDE-SOTO, G. La demostración matemática: significado, tipos, funciones atribuidas y relevancia en el conocimiento profesional de los profesores de matemáticas. *Uniciencia*, [online], v. 33, n. 2, p. 55-75, 2019. DOI: <http://dx.doi.org/10.15359/ru.33-2.5>.
- ALLAHBAKHSI, M., BEHN, A., HENAO, D. LEIVA, N.; MENARES, R. *Orientaciones para la construcción y la escritura de una demostración*. Proyecto Ciencia 2030. Pontificia Universidad Católica de Chile, 2022
- ALVARADO, A. & GONZÁLEZ, M.T. La implicación lógica en el proceso de demostración matemática: un estudio de caso. *Enseñanza de las ciencias*, v. 28, n. 1, p. 73-84, 2009.
- BUSTOS, A; ZUBIETA, G. Desarrollo y cambios en las maneras de justificar matemáticamente de estudiantes cuando trabajan en un ambiente sociocultural. *Enseñanza de las ciencias*, v. 37, n. 3, p. 129-148, 2019. DOI: <https://doi.org/10.5565/rev/ensciencias.2506>
- CAMACHO, V., SÁNCHEZ, J.J. & ZUBIETA, G. Los estudiantes de ciencias, ¿pueden reconocer los argumentos lógicos involucrados en una demostración? *Enseñanza de las ciencias*, v. 32, n. 1, p. 117-138, 2014. DOI: <https://doi.org/10.5565/rev/ensciencias.983>
- CHRISTIE, F. & DEREWIANKA, B. *School discourse: Learning to Write Across the Years of Schooling*. New York: Continuum Discourse, 2008.
- COCKING, R. R. & MESTRE, J. P. *Linguistic and cultural influences on learning mathematics*. Taylor & Francis Group, 1988

- COFFIN, C. Learning to write history. The role of causality. *Written Communication*, v. 21, n. 3, p. 261-289, 2004. DOI: <https://doi.org/10.1177/0741088304265476>
- DORAN, Y. J. *The Discourse of Physics: Building Knowledge through Language, Mathematics and Image*, London: Routledge, 2018a.
- DORAN, Y. J. Intrinsic Functionality of Mathematics, Metafunctions in Systemic Functional Semiotics, *Semiotica*, v. 225, p. 457-87, 2018b. DOI: <https://doi.org/10.1515/sem-2017-0004>
- DORAN, Y. J. Academic Formalisms: Toward a Semiotic Typology. In: MARTIN, J.R.; DORAN, Y.; FIGUEREDO, G. (eds). *Systemic Functional Language Description: Making Meaning Matter*. London: Routledge, 2020. p. 331-358.
- DORAN, Y. J. Semiotic Description: Grappling with Mathematics. In: CALDWELL, J., KNOX, J., & MARTIN, J. R. (eds.). *Applicable linguistics and social semiotics*. Bloomsbury Academic, 2022. p. 341-354.
- DORAN, Y. J. & MARTIN, J.R. Field relations: Understanding scientific explanations. In: K. MATON, K.; MARTIN, J.R.; DORAN, Y. J. (eds). *Studying Science: Knowledge, Language, Pedagogy*. London: Routledge, 2021.
- EGGINS, S. *An Introduction to Systemic Functional Linguistics*. Londres: Bloomsbury, 2004.
- FIALLO, J., CAMARGO, L. & GUTIERREZ, A. Acerca de la enseñanza y el aprendizaje de la demostración en matemáticas. *Revista Integración. Escuela de Matemáticas*, v. 31, n. 2, p. 181-205, 2013.
- HALLIDAY, M.A.K. *An Introduction to Functional Grammar*, 2 ed. Londres: Edward Arnold, 1994
- HALLIDAY, M.A.K. & MARTIN, J. R. *Writing science: Literacy and discursive power*. Londres: The Falmer Press, 1993.
- HALLIDAY, M.A.K. & C.M.I.M. MATTHIESSEN. *Halliday's Introduction to Functional Grammar*. 4. ed. Nueva York: Routledge, 2014.
- HALLIDAY, M. A. K. & HASAN, R. *Cohesion in English*. London: Longman, 1976.
- HAO, J. *Analysing Scientific Discourse: A framework for exploring knowledge building in biology from a systemic functional linguistic perspective*. New York: Routledge, 2020.
- HEYVAERT, L. Nominalization as Grammatical Metaphor: On the Need for a Radically Systemic and Metafunctional Approach. In: SIMON-VANDENBERGEN, A. M.; TAVERNIERS, M.; RAVELLI, L. (eds). *Grammatical Metaphor: Views from Systemic Functional Linguistics*, Ámsterdam, Filadelfia: John Benjamins, 2003. p. 65-99.
- KATZ, B.P., THOREN, E. & HERNÁNDEZ, V. Why Should that Convince Me?: Teaching Toulmin Analysis Across the Curriculum. *PRIMUS*, v. 33, n. 2, p. 285-313, 2023. DOI: <https://doi.org/10.1080/10511970.2022.2068093>
- LEIVA, N. Conexiones causales en español: un recurso semántico-discursivo para explicar el pasado reciente en la Historia escolar. *Estudios Filológicos*, n.69, p.135-161, 2022. DOI: <http://dx.doi.org/10.4067/S0071-17132022000100135>
- LEIVA, N. & OTEÍZA, T. Causalidad y posicionamientos en el discurso de la historia escolar en español. *Íkala, Revista de Lenguaje y Cultura*, v. 28, n.3, p.1-19, 2023. DOI: <https://doi.org/10.17533/udea.ikala.352539>.

- LEW, K. & MEJÍA-RAMOS, J. P. Linguistic conventions of mathematical proof writing across pedagogical contexts. *Educational Studies in Mathematics*, v. 103, n. 1, p. 43-62, 2020. DOI: <https://doi.org/10.1007/s10649-019-09915-5>
- MARTIN, J.R. *English text: system and structure*. Amsterdam: Benjamins, 1992
- MARTIN, J.R. Life as a noun: Arresting the universe in science and humanities. In: HALLIDAY, M. A. K.; MARTIN, J. R. (eds). *Writing science: Literacy and discursive power*. London: The Falmer Press, 1993. p. 242–293.
- MARTIN, J.R. Making History: Grammar for interpretation. In: MARTIN, J. R.; WODAK, R. (eds.). *Re-reading the past: Critical and functional perspectives on time and value*. Amsterdam: Benjamins, 2004, p. 20–56.
- MARTIN, J.R. *Genre and Field: Social Processes And Knowledge Structures in Systemic Functional Semiotics*. The 33rd International Systemic Functional Congress, 2006, *Proceedings...*
- MARTIN, J.R. *Ideational semiosis: a tri-stratal perspective on grammatical metaphor*. D.E.L.T.A., São Paulo, v. 26, n.2, p. 1-27, 2020. DOI: <https://doi.org/10.1590/1678-460X2020360304>.
- MARTIN, J.R. & WHITE, P.R.R. *The Language of Evaluation*. Londres: Palgrave, 2005.
- MARTIN, J.R. & ROSE, D. *Working With Discourse. Meaning Beyond the Clause*. London: Continuum, 2007.
- MARTIN, J.R. & ROSE, D. *Genre Relations. Mapping Culture*. London: Equinox Publishing, 2008.
- MARTIN, J. R., QUIROZ, B. & WANG, P. Theme. In: *Systemic Functional Grammar: A text-based description of English, Spanish and Chinese*. Cambridge: Cambridge University Press, 2023. p. 304-369.
- MARTÍNEZ, A. *La demostración en matemática. Una aproximación epistemológica y didáctica*. Quinto Simposio de la Sociedad Española de Investigación en Educación Matemática, 2001.
- MARTÍNEZ L. Algunas apreciaciones acerca del concepto crítico de demostración. *Logos. Anales del Seminario de Metafísica*, vol. 55, n. 1, p. 109-124, 2022. DOI: <https://doi.org/10.5209/asem.76153>
- MOYANO, E.I. El sistema de Tema en español: una mirada discursiva sobre una cuestión controversial. In: GHIO, E; FERNÁNDEZ, M.D. (eds.). *El discurso en español y portugués: Estudios desde una perspectiva sistémico-funcional*, Santa fe, Universidad Nacional del Litoral, 2010, p. 39-87
- MOYANO, E.I. La función de Tema en español: sus medios de realización desde la perspectiva trinocular de la Lingüística Sistémico Funcional. *Revista Signos*, v. 54, n. 106, p. 487-517, 2021. DOI: [10.4067/S0718-09342021000200487](https://doi.org/10.4067/S0718-09342021000200487).
- O'HALLORAN, K. Towards a systemic functional analysis of multisemiotic mathematics texts'. *Semiotica*, v. 124, n.1/2, p. 1-29, 1999. DOI: <https://doi.org/10.1515/semi.1999.124.1-2.1>
- O'HALLORAN, K. *Mathematical discourse. Language, symbolism and visual images*. London: Continuum, 2005.
- O'HALLORAN, K. The language of learning mathematics: A multimodal perspective. *Journal of Mathematical Behavior*, v.40, p. 63–74, 2015. DOI: <https://doi.org/10.1016/j.jmathb.2014.09.002>
- PIMM, D. *Speaking mathematically: Communication in mathematics classrooms*. London: Routledge, 1987.
- QUIROZ, B. & MARTIN, J. R. Perfil sistémico-funcional del grupo nominal en español: estructura, funciones discursivas básicas y organización sistémica. *Estudios Filológicos*, v. 68, p. 123–151, 2021. DOI: <https://doi.org/10.4067/S0071-17132021000200123>

- SCHLEPPGRELL, M. J. The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly*, v. 23, n. 2, p. 139–159, 2007. DOI: <https://doi.org/10.1080/10573560601158461>
- SEGERBY, C. *Supporting mathematical reasoning through reading and writing in mathematics: making the implicit explicit*. Holmbergs, Malmö University, 2017.
- SPANOS, G., RHODES, N., DALE, T. & CRANDALL, J. Linguistics Features of mathematical Problem Solving. In: COCKING, R. R.; MESTRE, J. P. (eds.). *Linguistic and cultural influences on learning mathematics*. Taylor & Francis Group, 1988. p. 221–240.
- SUA FLORES, C. Saber suficiente no es suficiente: comportamientos metacognitivos al resolver problemas de demostración con el apoyo de la geometría dinámica. *Tecne, episteme y didaxis: revista de la Facultad de Ciencia y Tecnología, Universidad Pedagógica Nacional*, n. 45, p. 121-142, 2019. DOI: <https://doi.org/10.17227/ted.num45-9838>
- SUNDSTROM, T. *Mathematical Reasoning: Writing and Proof*. Pearson Education, 2021.
- URHAN, S. & ZENGİN, Y. Investigating University Students' Argumentations and Proofs Using Dynamic Mathematics Software in Collaborative Learning, Debate, and Self-reflection Stages. *International Journal of Research in Undergraduate Mathematics Education*, v.10, n.2, p. 380–407, 2024. DOI: <https://doi.org/10.1007/s40753-022-00207-7>
- WILSON, J. A Primer on Mathematical Proof. <https://dept.math.lsa.umich.edu/~jch/w/2015Math110Material/PrimerOnProof-Math110.pdfs/a>

Comissão científica

Aderlanne Pereira Ferraz (UFMG, Belo Horizonte/MG, Brasil)
Alessandro Panunzi (Unifi, Florença, Itália)
Alina M. S. M. Villalva (ULisboa, Lisboa, Portugal)
Aline Alves Ferreira (UCSB, Santa Barbara/CA, Estados Unidos)
Ana Lúcia de Paula Müller (USP, São Paulo/SP, Brasil)
Ana Maria Carvalho (UA, Tucson/AZ, Estados Unidos)
Ana Paula Scher (USP, São Paulo/SP, Brasil)
Anabela Rato (U of T, Toronto/ON, Canadá)
Aparecida de Araújo Oliveira (UFV, Viçosa/MG, Brasil)
Aquiles Tescari Neto (UNICAMP, Campinas/SP, Brasil)
Augusto Soares da Silva (UCP, Braga, Portugal)
Beth Brait (PUC-SP/USP, São Paulo/SP, Brasil)
Bruno Neves Rati de Melo Rocha (UFMG, Belo Horizonte/MG, Brasil)
Carmen Lucia Barreto Matzenauer (UCPEL, Pelotas/RS, Brasil)
Celso Ferrarezi (UNIFAL, Alfenas/MG, Brasil)
César Nardelli Cambraia (UFMG, Belo Horizonte/MG, Brasil)
Cristina Name (UFJF, Juiz de Fora/MG, Brasil)
Charlotte C. Galves (UNICAMP, Campinas/SP, Brasil)
Deise Prina Dutra (UFMG, Belo Horizonte/MG, Brasil)
Diana Luz Pessoa de Barros (USP/UPM, São Paulo/SP, Brasil)
Edwiges Morato (UNICAMP, Campinas/SP, Brasil)
Emília Mendes Lopes (UFMG, Belo Horizonte/MG, Brasil)
Esmeralda V. Negrão (USP, São Paulo/SP, Brasil)
Flávia Azeredo Cerqueira (JHU, Baltimore/MD, Estados Unidos)
Gabriel de Avila Othero (UFRGS, Porto Alegre/RS, Brasil)
Gerardo Augusto Lorenzino (TU, Filadélfia/PA, Estados Unidos)
Glaucia Muniz Proença de Lara (UFMG, Belo Horizonte/MG, Brasil)
Hanna Batoréo (UAb, Lisboa, Portugal)
Heliana Ribeiro de Mello (UFMG, Belo Horizonte/MG, Brasil)
Heronides Moura (UFSC, Florianópolis/SC, Brasil)
Hilario Bohn (UCPEL, Pelotas/RS, Brasil)
Hugo Mari (PUC-Minas, Belo Horizonte/MG, Brasil)

Ida Lucia Machado (UFMG, Belo Horizonte/MG, Brasil)
Ieda Maria Alves (USP, São Paulo/SP, Brasil)
Ivã Carlos Lopes (USP, São Paulo/SP, Brasil)
Jairo Nunes (USP, São Paulo/SP, Brasil)
Jairo Venício Carvalhais Oliveira (UFMG, Belo Horizonte/MG, Brasil)
Jean Cristtus Portela (UNESP-Araraquara, Araraquara/SP, Brasil)
João Antônio de Moraes (UFRJ, Rio de Janeiro/ RJ, Brasil)
João Miguel Marques da Costa (Universidade Nova da Lisboa, Lisboa, Portugal)
João Queiroz (UFJF, Juiz de Fora/MG, Brasil)
José Magalhaes (UFU, Uberlândia/MG, Brasil)
João Saramago (Universidade de Lisboa, Lisboa, Portugal)
José Borges Neto (UFPR, Curitiba/PR, Brasil)
Kanavillil Rajagopalan (UNICAMP, Campinas/SP, Brasil)
Laura Alvarez Lopez (Universidade de Estocolmo, Stockholm, Suécia)
Leo Wetzels (Free Univ. of Amsterdam, Amsterdã, Holanda)
Laurent Filliettaz (Université de Genève, Genebra, Suíça)
Leonel Figueiredo de Alencar (UFC, Fortaleza/CE, Brasil)
Livia Oushiro (UNICAMP, Campinas/SP, Brasil)
Lodenir Becker Karnopp (UFRGS, Porto Alegre/RS, Brasil)
Lorenzo Teixeira Vitral (UFMG, Belo Horizonte/MG, Brasil)
Luiz Amaral (UMass Amherst, Amherst/MA, Estados Unidos)
Luiz Carlos Cagliari (UNESP, São Paulo/SP, Brasil)
Marcelo Barra Ferreira (USP, São Paulo/SP, Brasil)
Marcia Cançado (UFMG, Belo Horizonte/MG, Brasil)
Márcio Leitão (UFPB, João Pessoa/PB, Brasil)
Marcus Maia (UFRJ, Rio de Janeiro/RJ, Brasil)
Maria Bernadete Marques Abaurre (UNICAMP, Campinas/SP, Brasil)
Maria Cecília Camargo Magalhães (PUC-SP, São Paulo/SP, Brasil)
Maria Cecília Magalhães Mollica (UFRJ, Rio de Janeiro/RJ, Brasil)
Maria Cândida Trindade Costa de Seabra (UFMG, Belo Horizonte/MG, Brasil)
Maria Cristina Figueiredo Silva (UFPR, Curitiba/PR, Brasil)
Maria Luíza Braga (PUC/RJ, Rio de Janeiro/RJ, Brasil)
Maria Marta P. Scherre (UNB, Brasília/DF, Brasil)
Micheline Mattedi Tomazi (UFES, Vitória/ES, Brasil)
Miguel Oliveira, Jr. (UFAL, Maceió, Alagoas, Brasil)
Monica Santos de Souza Melo (UFV, Viçosa/MG, Brasil)
Patricia Matos Amaral (UI, Bloomington/IN, Estados Unidos)
Paulo Roberto Gonçalves Segundo (USP, São Paulo/SP, Brasil)
Philippe Martin (Université Paris 7, Paris, França)
Rafael Nonato (Museu Nacional-UFRJ, Rio de Janeiro/RJ, Brasil)
Raquel Meister Ko. Freitag (UFS, Aracaju/SE, Brasil)
Renato Miguel Basso (UFSCar, São Carlos, SP, Brasil).
Roberto de Almeida (Concordia University, Montreal/QC, Canadá)
Ronice Müller de Quadros (UFSC, Florianópolis/SC, Brasil)
Ronald Beline (USP, São Paulo/SP, Brasil)
Rove Chishman (UNISINOS, São Leopoldo/RS, Brasil)
Sanderléia Longhin-Thomazi (UNESP, São Paulo/SP, Brasil)
Sergio de Moura Menuzzi (UFRGS, Porto Alegre/RS, Brasil)
Seung- Hwa Lee (UFMG, Belo Horizonte/MG, Brasil)

Sírio Possenti (UNICAMP, Campinas/SP, Brasil)
Suzi Lima (U of T / UFRJ, Toronto/ON - Rio de Janeiro/RJ, Brasil)
Thais Cristofaro Alves da Silva (UFMG, Belo Horizonte/MG, Brasil)
Tommaso Raso (UFMG, Belo Horizonte/MG-Brasil)
Tony Berber Sardinha (PUC-SP, São Paulo/SP, Brasil)
Ubiratã Kickhöfel Alves (UFRGS, Porto Alegre/RS, Brasil)
Vander Viana (University of Stirling, Stirling/Sld, Reino Unido)
Vanise Gomes de Medeiros (UFF, Niterói/RJ, Brasil)
Vera Lucia Lopes Cristovao (UEL, Londrina/PR, Brasil)
Vera Menezes (UFMG, Belo Horizonte/MG, Brasil)
Vilson José Leffa (UCPel, Pelotas/RS, Brasil)